

Double differential charm and beauty production cross section measurements at 5 TeV in CMS at the LHC

Dissertation

zur Erlangung des Doktorgrades
an der Fakultät für Mathematik, Informatik
und Naturwissenschaften Fachbereich Physik der
Universität Hamburg

vorgelegt von

Josry Metwally

This dissertation is submitted for the degree of
Doctor rerum naturalium

Hamburg
2022



Gutachter der Dissertation:

Dr. Achim Geiser

Prof. Dr. Elisabetta Gallo

Zusammensetzung der Prüfungskommission:

Dr. Achim Geiser

Prof. Dr. Elisabetta Gallo

Prof. Dr. Freya Blekman

Prof. Dr. Dieter Horns

Dr. Markus Diehl

Vorsitzende des Prüfungsausschusses:

Prof. Dr. Dieter Horns

Datum der Disputation:

23. Juni 2022

Vorsitzender Fach-Promotionsausschuss Physik:

Prof. Dr. Wolfgang J. Parak

Leiter des Fachbereichs Physik:

Prof. Dr. Günter H. W. Sigl

Dekan der Fakultät MIN:

Prof. Dr. Heinrich Graener

Eidesstattliche Versicherung / Declaration on oath

Hiermit versichere ich an Eides statt, die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt zu haben.

Hamburg, 31. May 2022

A handwritten signature in blue ink, consisting of a stylized 'J' and 'M'.

Josry Metwally

Abstract

Double differential cross sections for the production of prompt D^* mesons, and D^* mesons from b hadron decays (non-prompt) are measured through the decays $D^* \rightarrow D^0 \pi_s \rightarrow K \pi \pi_s$ and $B \rightarrow D^* X \rightarrow D^0 \pi_s X \rightarrow K \pi \pi_s X$ in pp collisions in CMS at a center of mass energy of 5 TeV in the full phase space complementary to LHCb. The separation of prompt from non-prompt D^* mesons near the production threshold is performed in the kinematic range $p_T > 1.5$ GeV and $|y| < 2$. This procedure and the resulting cross sections in the accessible phase space including a comparison with QCD theory predictions and comparisons with other experiments are presented within this thesis.

The procedure including the results of this thesis opens the opportunity for the determination of cross sections for inclusive charm and beauty production at the LHC with different center-of-mass energies down to very low transverse momentum, and the comparison with QCD predictions in next-to-leading order of perturbation theory.

Zusammenfassung

Doppelt differenzielle Wirkungsquerschnitte für die Produktion prompter D^* -Mesonen und D^* -Mesonen aus b-Hadronzerfällen (nicht-prompt) werden durch die Zerfälle $D^* \rightarrow D^0 \pi_s \rightarrow K \pi \pi_s$ und $B \rightarrow D^* X \rightarrow D^0 \pi_s X \rightarrow K \pi \pi_s X$ in pp-Kollisionen in CMS bei einer Schwerpunktsenergie von 5 TeV im gesamten Phasenraum komplementär zu LHCb gemessen. Die Separation von prompten D^* -Mesonen und D^* -Mesonen aus b-Hadronzerfällen erfolgt nahe der Produktionsschwelle bei $p_T > 1.5$ GeV und im Pseudorapiditätsbereich $|y| < 2$. In dieser Arbeit wird dieses Verfahren und die daraus resultierenden Wirkungsquerschnitte im zugänglichen Phasenraum vorgestellt, einschließlich eines Vergleichs mit Vorhersagen der QCD-Theorie und Vergleichen mit anderen Experimenten.

Das Verfahren unter Berücksichtigung der Ergebnisse dieser Arbeit eröffnet die Möglichkeit zur Bestimmung von Wirkungsquerschnitten für inklusive Charm- und Beauty-Quark-Produktion am LHC mit unterschiedlichen Schwerpunktsenergien bis hinunter zu sehr niedrigen Transversalimpulsen und ermöglicht den Vergleich mit QCD-Vorhersagen in nächstführender Ordnung in der Störungstheorie.

Table of contents

1	Introduction	1
2	The Standard Model of Particle Physics	5
2.1	Elementary Particles and their Interactions	6
2.1.1	Fermions	7
2.1.2	Bosons	7
2.2	Theoretical Formulation	7
2.2.1	Feynman diagrams	8
2.2.2	Gauge theories	9
2.2.3	Electroweak interaction	10
2.2.4	Lagrangian of QCD	13
2.2.5	The Higgs sector	15
2.3	QCD Cross Section in Hadron Hadron Collisions	18
2.3.1	Parton model and PDFs	20
2.3.2	Hard process and factorization	22
2.3.3	Parton shower and hadronization	24
	References	27
3	Physics in CMS with the LHC at CERN	31
3.1	CERN	32
3.1.1	The LHC	32
3.2	The Compact Muon Solenoid Detector	33
3.2.1	The coordinate system of CMS	36
3.2.2	Magnet	36
3.2.3	Muon system	37
3.2.4	Hadronic calorimeter	38
3.2.5	Electromagnetic calorimeter	38
3.3	The CMS Tracker and its Alignment	39

3.3.1	The silicon tracker	39
3.3.2	Tracker alignment	41
	References	47
4	Event Reconstruction and Data Acquisition	49
4.1	Tracks and Vertices	50
4.2	Muons and Electrons	51
4.3	Particle Flow	52
4.4	Jets	53
4.4.1	Clustering	53
4.4.2	b-jet identification	54
4.5	Missing Transverse Energy	54
4.6	Acquisition and Analysis of Data	55
4.6.1	Trigger system	56
4.6.2	Data formats and analysis	57
4.7	MC and Detector Simulations	58
	References	61
5	Physics of Heavy Quark Flavor Production	65
5.1	Heavy Flavor Production and Decays in pp Collisions	65
5.2	Theory Predictions for Heavy Flavor Production	68
5.3	Existing Measurements of Charm and Beauty Production	71
	References	81
6	Charm and Beauty Cross Sections	85
6.1	Data and Monte Carlo Samples	86
6.1.1	Effective integrated luminosity for pp minimum bias data at 5 TeV	88
6.1.2	Effective luminosity for MC samples	88
6.2	Event Selection and D^* Reconstruction	89
6.3	Background Estimation and Signal Extraction	94
6.4	Determination of the Efficiency	96
6.5	Double Differential Cross Sections for Inclusive D^* production	99
6.6	Separation of Charm and Beauty in D Meson Final States	103
6.7	A Different Ansatz for the Charm Beauty Separation	112
6.8	Double Differential Prompt and Non-prompt D^* Cross Sections	114
6.9	Systematic Uncertainties	122
	References	123

7 Conclusion and Outlook	125
Appendix A Track Based Alignment	127
A.1 Weak Modes	127
A.2 Testing the Impact on Reconstructed Secondary Vertices due to the Improve- ment of the Alignment	128
References	131
Appendix B Measurement of inclusive D^* Cross Sections in Bins of p_T and Rapid- ity	133
B.1 Effective Integrated Luminosity for pp Minimum Bias Data at 5 TeV	133
B.2 Signal Extraction for D^* in Bins of p_T and Rapidity	134
References	145
Appendix C Measurement of Prompt and Non-prompt D^* Cross Sections in Bins of p_T and Rapidity	147
C.1 Signal Extraction for D^* in Bins of p_T and Rapidity	147
C.2 Signal Extraction at Truth Level for Prompt and Non-prompt D^* in Bins of p_T and Rapidity	158
C.3 Separation of Charm and Beauty in D^* Final States in Bins of p_T and Rapidity	178
C.4 Investigate of Correlations - A Different Ansatz for the Charm Beauty Sepa- ration	198
List of figures	201
List of tables	213

Chapter 1

Introduction

The description of fundamental interactions between elementary particles is one goal of particle physics. The Standard Model of particle physics (SM) [1–3] describes the particles and their interactions. So far, the predictions of the SM are in good agreement with experimental measurements. Since the SM is a gauge theory with left-right symmetry breaking, particles are a priori massless. However, it is well known that nearly all elementary particles carry mass, for instance, the electron. In 2012, the CMS¹ and ATLAS² Collaborations discovered a boson at the Large Hadron Collider (LHC) operated by the European Organization for Nuclear Research (CERN), which has properties consistent with to the SM Higgs boson [4, 5]. The problem of massive particles in the SM is solved by the Brout-Englert-Higgs mechanism [6–8], which leads additionally to the prediction of such a further particle, the scalar Higgs boson that couples to all massive particles of the SM.

Higgs couplings, for instance to heavy quarks, can be tested directly with the recorded data by CMS at the LHC. But they can also be tested indirectly by for instance the measurement of the cross sections for the production of heavy quarks at different center of mass (c.m.) energies, which in turn can be used for a measurement of the quark mass and converted to the corresponding Higgs boson quark coupling. Measurements at different c.m. energies exist for top pair production, as shown in Fig. 1.1. In the remaining heavy quark sector (charm and beauty), other experiments such as ATLAS and ALICE³ covered only small fractions of the available phase space, while the LHCb⁴ experiment almost fully covered the forward rapidity (y) region, $2.0 < y < 4.5$. The measurement of the charm and beauty production cross sections of this thesis are used for a project of the QCD group at DESY⁵ measuring the

¹Compact Muon Solenoid

²A Toroidal LHC ApparatuS

³A Large Ion Collider Experiment

⁴Large Hadron Collider beauty

⁵Deutsches Elektronen-Synchrotron

total charm and beauty production cross sections at different c.m. energies at the LHC in proton-proton (pp) collisions.

In this work, cross sections of prompt D^* mesons and D^* mesons from b hadron decays are measured double differentially through the decays $D^* \rightarrow D^0 \pi_s \rightarrow K \pi \pi_s$ and $B \rightarrow D^* X \rightarrow D^0 \pi_s X \rightarrow K \pi \pi_s X$. For this purpose, a charm/beauty separation is performed to extract the individual cross sections for prompt D^* (charm) and non-prompt D^* (beauty), respectively. The cross sections for charm and beauty production are measured as a function of the transverse momentum p_T and the rapidity of the corresponding D^* in the full accessible phase space of CMS complementary to LHCb at $\sqrt{s} = 5$ TeV in this work.

This thesis starts with a brief theoretical background underlying this work in Chapter 2. The LHC and the CMS detector, which records the collision data used for this work, are described in Chapter 3. For an analysis, an event consisting of several objects needs to be reconstructed, which is explained in Chapter 4. A description of the physics of charm and beauty production in pp collisions, the used theory tools for this work and the corresponding relevant existing results are given in Chapter 5. The analysis strategy, the separation of prompt D^* mesons and D^* mesons from b hadron decays near the production threshold and the resulting cross sections are explained in Chapter 6. A summary of this work including an outlook is given in Chapter 7.

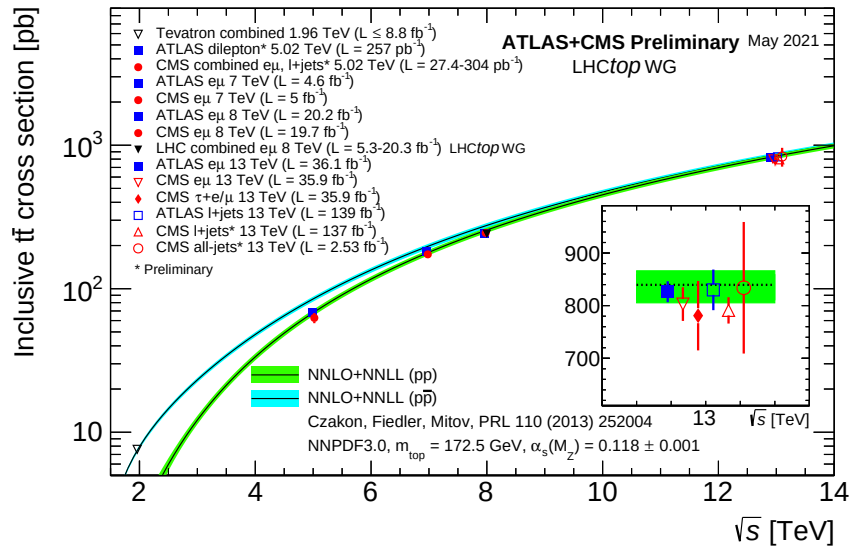


Fig. 1.1 Top pair production as a function of center of mass energy for $m_t = 172.5 \text{ GeV}$. It includes the measurements of LHC and Tevatron which are compared to the NNLO QCD calculation complemented with NNLL resummation (top++2.0) [9]. The theory band arises due to the uncertainties coming from renormalisation and factorisation scale, parton density functions and the strong coupling. The different measurements are slightly offset in $\sqrt{s}$ for better visibility. Taken from the LHCTopWG [10].

Chapter 2

The Standard Model of Particle Physics

Contents

2.1	Elementary Particles and their Interactions	6
2.1.1	Fermions	7
2.1.2	Bosons	7
2.2	Theoretical Formulation	7
2.2.1	Feynman diagrams	8
2.2.2	Gauge theories	9
2.2.3	Electroweak interaction	10
2.2.4	Lagrangian of QCD	13
2.2.5	The Higgs sector	15
2.3	QCD Cross Section in Hadron Hadron Collisions	18
2.3.1	Parton model and PDFs	20
2.3.2	Hard process and factorization	22
2.3.3	Parton shower and hadronization	24
	References	27

This chapter provides the theoretical background underlying this work. It starts with an overview of elementary particles and their properties. Together with their interactions these are mathematically described in the Lagrangian of the Standard Model of particle physics (SM). Furthermore it will be explained how the SM can be used to make theoretical predictions that can be compared to experimental measurements and, therefore, used to validate the SM and to extract its parameters.

Since the focus is to give an overview of the theory of particle physics and to put this thesis in the right context, some aspects are discussed only briefly. More detailed descriptions can be found in References [11–15].

2.1 Elementary Particles and their Interactions

The SM describes the properties and interactions of all known elementary particles. Starting with the *electron* that was already discovered in 1897 [16] and ending with the *Higgs boson* that was finally discovered in 2012 [4, 5]. Electromagnetic interactions like the photoelectric effect, weak interactions like the β -decay or even interactions responsible for binding nuclei due to the strong force are embedded in the SM. A collection of all elementary particles can be found in Fig. 2.1. In general they can be divided into two groups, *fermions* and *bosons*, with distinct properties.

Standard Model of Elementary Particles							
three generations of matter (elementary fermions)			three generations of antimatter (elementary antifermions)			interactions / force carriers (elementary bosons)	
I	II	III	I	II	III		
mass charge spin	$\approx 2.2 \text{ MeV}/c^2$ $\frac{2}{3}$ $\frac{1}{2}$	$\approx 1.28 \text{ GeV}/c^2$ $\frac{2}{3}$ $\frac{1}{2}$	$\approx 173.1 \text{ GeV}/c^2$ $\frac{2}{3}$ $\frac{1}{2}$	$\approx 2.2 \text{ MeV}/c^2$ $-\frac{2}{3}$ $\frac{1}{2}$	$\approx 1.28 \text{ GeV}/c^2$ $-\frac{2}{3}$ $\frac{1}{2}$	$\approx 173.1 \text{ GeV}/c^2$ $-\frac{2}{3}$ $\frac{1}{2}$	0 0 1
QUARKS	u up	c charm	t top	$\bar{u}$ antiup	$\bar{c}$ anticharm	$\bar{t}$ antitop	g gluon
	d down	s strange	b bottom	$\bar{d}$ antidown	$\bar{s}$ antistrange	$\bar{b}$ antibottom	γ photon
	$\approx 4.7 \text{ MeV}/c^2$ $-\frac{1}{3}$ $\frac{1}{2}$	$\approx 96 \text{ MeV}/c^2$ $-\frac{1}{3}$ $\frac{1}{2}$	$\approx 4.18 \text{ GeV}/c^2$ $-\frac{1}{3}$ $\frac{1}{2}$	$\approx 4.7 \text{ MeV}/c^2$ $\frac{1}{3}$ $\frac{1}{2}$	$\approx 96 \text{ MeV}/c^2$ $\frac{1}{3}$ $\frac{1}{2}$	$\approx 4.18 \text{ GeV}/c^2$ $\frac{1}{3}$ $\frac{1}{2}$	0 0 1
LEPTONS	e electron	μ muon	τ tau	e^+ positron	μ^+ antimuon	τ^+ antitau	Z Z ⁰ boson
	$\approx 0.511 \text{ MeV}/c^2$ -1 $\frac{1}{2}$	$\approx 105.66 \text{ MeV}/c^2$ -1 $\frac{1}{2}$	$\approx 1.7768 \text{ GeV}/c^2$ -1 $\frac{1}{2}$	$\approx 0.511 \text{ MeV}/c^2$ 1 $\frac{1}{2}$	$\approx 105.66 \text{ MeV}/c^2$ 1 $\frac{1}{2}$	$\approx 1.7768 \text{ GeV}/c^2$ 1 $\frac{1}{2}$	$\approx 91.19 \text{ GeV}/c^2$ 0 1
	$< 2.2 \text{ eV}/c^2$ 0 $\frac{1}{2}$	$< 0.17 \text{ MeV}/c^2$ 0 $\frac{1}{2}$	$< 18.2 \text{ MeV}/c^2$ 0 $\frac{1}{2}$	$< 2.2 \text{ eV}/c^2$ 0 $\frac{1}{2}$	$< 0.17 \text{ MeV}/c^2$ 0 $\frac{1}{2}$	$< 18.2 \text{ MeV}/c^2$ 0 $\frac{1}{2}$	$\approx 80.39 \text{ GeV}/c^2$ 1 1
	ν_e electron neutrino	ν_μ muon neutrino	ν_τ tau neutrino	$\bar{\nu}_e$ electron antineutrino	$\bar{\nu}_\mu$ muon antineutrino	$\bar{\nu}_\tau$ tau antineutrino	W⁺ W ⁺ boson
							W⁻ W ⁻ boson
							$\approx 124.97 \text{ GeV}/c^2$ 0 0
							H higgs
							SCALAR BOSONS
							GAUGE BOSONS VECTOR BOSONS

Fig. 2.1 Overview of the elementary particles of the Standard Model of particle physics and their anti-particles. The quarks are marked purple, the leptons green and the bosons red, except the Higgs boson which is shown in yellow. The anti-partners with their opposite quantum number are shown for all fermions too. This figure is taken from Reference [17] and modified.

2.1.1 Fermions

All particles that carry half integer spin are called fermions. The fermions are divided into leptons and quarks. The leptons are electron, muon, tau and their neutrino partners ($e, \nu_e, \mu, \nu_\mu, \tau, \nu_\tau$). They and their anti-partner are shown in Fig. 2.1. The electrical charge is -1 for charged leptons and $+1$ for their anti-partners. The neutrinos carry no electrical charge. The quarks are called up, down, charm, strange, top, and beauty (u, d, c, s, t, b). In addition to their electrical charge of $\frac{2}{3}$ or $-\frac{1}{3}$, they carry a color charge (red, blue or green). Their anti-partners carry the opposite electrical and color charge. All fermions appear in three generations marked with I-III in Fig. 2.1, with the first generation being the lightest in mass and the third the heaviest. Leptons appear as free particles while quarks do not. They do form hadrons, for instance, the proton. All fermions can interact with each other via the electroweak force but the strong force is only acting on colored particles.

2.1.2 Bosons

The interactions between fermions are happening via the exchange of bosons, the mediators of the different forces, also shown in Fig. 2.1. The mediators of the electroweak force are called photon γ , $W^\pm$ and Z boson. The photon and Z boson are electrically neutral but the $W^\pm$ boson carries an electric charge. The mediators of the strong force are called gluons. Those interact only with colored particles. Their color state consists of a combination of a color charge and anti-color charge, and therefore, gluon self-interaction is possible. All bosons carry a spin of 1 except the latest discovered one, the Higgs boson H with a spin of 0. The corresponding field is needed for the generation of mass terms for fermions and bosons without violating fundamental parts of the SM as described in Section 2.2.5. The W^+ is the anti-partner of the W^- boson, while all other bosons are their own anti-partner.

2.2 Theoretical Formulation

The mathematical formulation of the SM is the Lagrangian of the SM $\mathcal{L}_{SM}$. It comprises the fundamental interactions of elementary particles. The SM is a quantum field gauge theory (QFT) with the symmetry group $SU(3)_c \times SU(2)_L \times U(1)_Y$. The $SU(3)_c$ is a non-abelian group with 8 gauge bosons (gluons). The quantum number c represents the color charge of the quantum chromodynamics (QCD) formalism describing the strong force. The product $SU(2)_L \times U(1)_Y$ represents the unification of quantum electrodynamics (QED) and the weak force, called electroweak unification. The Hypercharge is Y and L indicates that the weak force distinguishes between left and right handed chirality, such that only the left handed

part enters the symmetry. Since the SM Lagrangian is forced to be gauge invariant, all 12 gauge bosons

$$\underbrace{\text{SU}(3)_c}_{8G_\mu^\lambda, \lambda=1,\dots,8} \times \underbrace{\text{SU}(2)_L}_{3W_\mu^a, a=1,2,3} \times \underbrace{\text{U}(1)_Y}_{B_\mu} \quad (2.1)$$

are forced to be massless, since their mass terms are not gauge invariant. Due to the Higgs-Sector of the SM, the $\text{SU}(2)_L \times \text{U}(1)_Y$ symmetry is spontaneously broken and the corresponding gauge bosons become massive except the photon, a linear combination of the W and B fields. Even for fermions mass terms can not enter the Lagrangian without violating gauge invariance, since they violate the asymmetry between left and right handed fermions. The massive Higgs boson with its interactions is formulated in the Higgs-Sector of the SM described in Section 2.2.5.

The SM Lagrangian before electroweak symmetry breaking (no explicit mass terms) can schematically be written as

$$\mathcal{L}_{SM} = \mathcal{L}_{QCD} + \mathcal{L}_{EW} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa}, \quad (2.2)$$

Here, $\mathcal{L}_{QCD}$ is the QCD Lagrangian, which describes the interaction of colored particles and $\mathcal{L}_{EW}$ is the electroweak Lagrangian, which describes the interaction of both leptons and quarks. The Higgs Lagrangian $\mathcal{L}_{Higgs}$ includes the kinematic Higgs term and the Higgs potential. Both are needed to obtain masses for the electroweak gauge bosons and the Higgs boson, respectively. The masses of fermions are obtained via the Yukawa terms $\mathcal{L}_{Yukawa}$ which are explained in Section 2.2.5. Before discussing gauge theory and the individual parts of the Lagrangian, respectively, a brief introduction to Feynman diagrams is given in the following.

2.2.1 Feynman diagrams

Since Feynman diagrams are used in this thesis to illustrate the studied processes, a very brief introduction is given in the following. Thanks to Richard Feynman [18], complicated mathematical expressions in $\mathcal{L}_{SM}$ representing interactions of elementary particles can also be represented easily as images, while at the same time corresponding to exactly defined integrals with the use of some specific rules. Because of this, such images are nowadays

known as Feynman diagrams. He considered how different particles and their vertices could be represented by drawing simply pictures of their interactions.

The particles are drawn as curved, curly, or continuous and dashed lines, depending on the type of particle. The interaction points (vertices) correspond to the crossing of these lines and are highlighted by a dot. Incoming and outgoing particles represent the initial and final state, respectively, and lines between the vertices correspond to a virtual particle in a Feynman diagram. For each additional vertex (coupling), the respective process is increasingly unlikely, since an additional vertex corresponds to a multiplicative factor proportional to the respective coupling strength of the particles. Depending on the field of the particles the coupling strength corresponds to α_s , in the case of strong interactions, or α_{ew} if particles interact via the electroweak force, or α_0 in the case of pure QED.

A much more detailed explanation of Feynman diagrams including their rules can be found in Reference [19]. Feynman diagrams in this thesis are drawn with JAXODRAW [20].

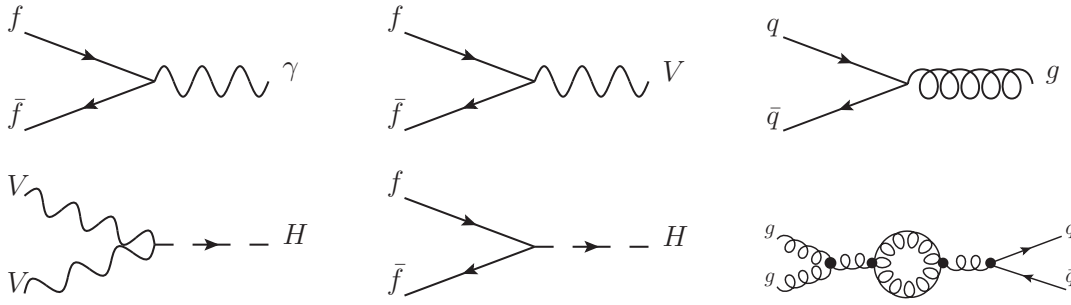


Fig. 2.2 Interaction between fermions and a photon (left), heavy vector boson V (middle) and gluon (right) are shown in the first row. Couplings between a scalar Higgs boson H with heavy vector bosons V or with a fermion anti-fermion pair are shown in the second row. A virtual correction is given with the last diagram, gluon loop correction. Vertices marked as dots are interactions between bosons and fermions or bosons and bosons. It is omitted to highlight those vertices here explicitly, except for the last diagram.

2.2.2 Gauge theories

The production processes for charm and beauty at the LHC studied within this thesis are dominated by the fusion of gluons, which are gauge bosons. Gauge bosons arise by forcing gauge invariance to a Lagrangian. To see how this works we consider QED as simple example. The kinematics of a relativistic fermion, e.g. an electron, with the field ψ can be described via the Dirac Lagrangian

$$\mathcal{L}_D = \bar{\psi} (i\gamma^\mu \partial_\mu - m) \psi, \quad (2.3)$$

where m represents the mass of the charged fermion and γ^μ the Dirac matrices. The Lagrangian is invariant under global $U(1)_{QED}$ transformations

$$\begin{aligned}\psi &\rightarrow \psi' = \exp(ie\theta)\psi(x), \\ \bar{\psi} &\rightarrow \bar{\psi}' = \exp(-ie\theta)\bar{\psi}(x).\end{aligned}\tag{2.4}$$

But under local transformation where $\theta = \theta(x)$ in Eq. 2.4 the invariance is broken. To regain invariance, thus an additional field has to be introduced such that $\mathcal{L}'_D = \mathcal{L}_D$ even under local transformations. We substitute the derivative ∂_μ in Eq. 2.3 with

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + ieA_\mu \tag{2.5}$$

where D_μ is the covariant derivative and A_μ the electromagnetic potential which can be identified with the photon if the vector field transforms like

$$A_\mu \rightarrow A'_\mu = A_\mu - \partial_\mu \theta(x), \tag{2.6}$$

and together with the transformation in Eq. 2.4, the Lagrangian in Eq. 2.3 is again invariant. Therefore, the Dirac Lagrangian in Eq. 2.3 becomes

$$\mathcal{L}_{QED} = \sum_{\text{fermions}} \bar{\psi}_f (iD - m) \psi_f - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \tag{2.7}$$

with the field strength tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{2.8}$$

The mass term for fermions for instance an electron is given in Eq. 2.7 by $m\bar{\psi}_f\psi_f$, while there is no mass for A_μ in the QED Lagrangian, reflecting that the mass of the photon is zero and gauge invariance is preserved.

2.2.3 Electroweak interaction

Beauty or charm quark decays arise in the SM only by the weak force. The successful description of the electromagnetic force as a gauge theory, leads to the formulation of weak interactions as a gauge theory, too [21, 22, 1, 23, 24]. For this a unified electroweak theory based on a non-abelian $SU(2)$ group is needed. The unification of the electromagnetism and

the weak force is described by the gauge theory formulated by Glashow-Weinberg-Salam¹ [2, 3, 21]. The non-abelian $SU(2)_L \times U(1)_Y$ symmetry group belongs to the electroweak theory where L symbols that the weak force distinguishes between left and right handed chirality and only the left handed part enters the symmetry. Y stands for the hypercharge described by the Gell-Mann-Nishijima formula [25, 26]

$$Y = 2(Q - T_3), \quad (2.9)$$

where Q stands for the electrical charge and T_3 for the third component of the weak isospin. Fields that transform under the electroweak gauge group are leptons and quarks. The left handed fermions ψ_L appear as isospin doublets of $SU(2)_L$

$$\ell_L^j = \begin{pmatrix} \nu_\ell^j \\ \ell^j \end{pmatrix}_L = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L, \quad (2.10)$$

$$q_L^j = \begin{pmatrix} u^j \\ d^j \end{pmatrix}_L = \begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L,$$

while the right handed fermions ψ_R are written as isospin singlets²

$$\begin{aligned} \ell_R = \ell_R^j &= e_R, \mu_R, \tau_R, \\ q_R = u_R^j, d_R^j &= u_R, c_R, t_R, d_R, s_R, b_R, \end{aligned} \quad (2.11)$$

that only transform under $U(1)$. The three generations are indicated in Eq. 2.11 by $j = 1, 2, 3$. The local transformation under $SU(2)_L \times U(1)_Y$ is given by

$$\begin{aligned} \psi_L &\rightarrow \psi'_L = e^{(i\theta(x)Y_L)} U_L \psi_L \\ \psi_R &\rightarrow \psi'_R = e^{(i\theta(x)Y_R)} \psi_R \end{aligned} \quad (2.12)$$

where $U_L = e^{i\vec{\alpha}\vec{\sigma}}$ is the one that only acts on the left handed fermions and σ corresponds to the Pauli matrices. The hypercharge values for the left handed fields can be derived from Eq. 2.9 and are listed for leptons and quarks in Tab. 2.1.

¹Meidner Sheldon Glashow, Steven Weinberg, and Abdus Salam got for their work the Nobel Prize in Physics in 1979.

²Assuming there are no right handed neutrinos in the SM.

Table 2.1 List of hypercharge values for left and right handed quarks and leptons. The index $j = 1, 2, 3$ stands for the three generations, T correspond to the isospin and T_3 to the third component of it. Note, there are no right handed neutrinos ν_R in the SM.

lepton	T	T_3	Y_L	quark	T	T_3	Y_L
$\nu_{L,\ell}^j$	$\frac{1}{2}$	$+\frac{1}{2}$	-1	u_L^j	$\frac{1}{2}$	$+\frac{1}{2}$	$\frac{1}{3}$
ℓ_L^j	$\frac{1}{2}$	$-\frac{1}{2}$	-1	d_L^j	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{3}$
lepton	T	T_3	Y_R	quark	T	T_3	Y_R
				u_R^j	0	0	$\frac{4}{3}$
ℓ_R^j	0	0	-2	d_R^j	0	0	$-\frac{2}{3}$

An additional ingredient to build a gauge invariant Lagrangian for the electroweak sector is the introduction of a covariant derivative that acts on ψ_L and ψ_R as

$$D_\mu^L = \partial_\mu - ig \sum_{a=1}^3 \tau^a W_\mu^a - i \frac{g'}{2} Y_L B_\mu, \quad D_\mu^R = \partial_\mu - i \frac{g'}{2} Y_R B_\mu. \quad (2.13)$$

Here, the gauge fields W_μ^a and B_μ and coupling constants g' and g belong to the $SU(2)_L$ and $U(1)_Y$ gauge group, respectively. The τ^a corresponds to $\sigma_a/2$. Starting with the Dirac Lagrangian in Eq. 2.3 for left and right handed fields, leads to

$$\mathcal{L}'_{EW} = \bar{\Psi}_L (i\gamma^\mu \partial_\mu) \Psi_L + \bar{\Psi}_R (i\gamma^\mu \partial_\mu) \Psi_R. \quad (2.14)$$

A substitution of ∂_μ with the covariant derivative in Eq. 2.13, the Lagrangian of the electroweak sector becomes

$$\begin{aligned} \mathcal{L}_{EW} &= \bar{\Psi}_L D_\mu^L \Psi_L + \bar{\Psi}_R D_\mu^R \Psi_R \\ &= \bar{\Psi}_L \left(i\gamma^\mu \partial_\mu - ig \sum_{a=1}^3 \tau^a W_\mu^a - i \frac{g'}{2} Y_L B_\mu \right) \Psi_L \\ &\quad + \bar{\Psi}_R \left(i\gamma^\mu \left(\partial_\mu - i \frac{g'}{2} Y_R B_\mu \right) \right) \Psi_R \\ &\quad - \frac{1}{4} B^{\mu\nu} B_{\mu\nu} - \frac{1}{4} W^{a\mu\nu} W_{\mu\nu}^a. \end{aligned} \quad (2.15)$$

Similar to QED, the field strength tensor for the gauge field B_μ is given by Eq. 2.8 by replacing A with B . In contrast for the gauge fields W_μ^a , the field strength tensor $W_{\mu\nu}^a$ is given

by

$$W_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g\epsilon^{abc}W_\mu^b W_\nu^c, \quad (2.16)$$

where ϵ^{abc} corresponds to the Levi-Civita tensor. Linear combinations of W_μ^a and B_μ lead to the physical fields $W_\mu^\pm, A_\mu, Z_\mu$. Since $SU(2)_L$ is handling the chirality of fermions differently, mass terms for them would break the gauge invariance. In addition, gauge theory is forcing the gauge bosons to be massless but it is experimentally observed that the gauge bosons from the weak force are massive. Thus, parts of the symmetry of $SU(2)_L \times U(1)_Y$ need to be spontaneously broken. How fermions and heavy vector bosons become massive is described in Section 2.2.5.

2.2.4 Lagrangian of QCD

Interactions between *colored* particles, quarks and gluons, are described by Quantum chromodynamics (QCD) constructed as a gauge theory with the corresponding gauge group $SU(3)_c$. The strong force is only acting on color charged particles. As in the case of QED, the symmetry is not broken and the mediators of the strong force remain massless. Similar to the electroweak sector, QCD is formulated as a non-abelian gauge theory leading to self-interactions between gluons. The QCD Lagrangian after electroweak symmetry breaking (i.e. with explicit quark mass terms) can schematically be written as

$$\mathcal{L}_{\text{QCD}} = \sum_{q=u,d,\dots} \bar{\psi}_{q,k} (i\gamma^\mu D_\mu - m_q)_{k,l} \psi_{q,l} - \frac{1}{4} G^{a\mu\nu} G_{\mu\nu}^a, \quad (2.17)$$

Here, the index q runs over all quark flavors, the indices k, l run over all three colors, γ^μ corresponds to the Dirac matrices explained in Section 2.2.2 and m_q to the quark mass. The covariant derivative and the field strength tensor are given by

$$D_\mu = \partial_\mu - ig_s \lambda^a G_\mu^a, \quad (2.18)$$

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a + g_s f^{abc} G_\mu^b G_\nu^c. \quad (2.19)$$

The eight Gell-Mann matrices λ^a in Eq. 2.18 correspond to the eight generators of the $SU(3)_c$ group. The eight gauge fields G_ν^a are the gluons, which are introduced by the covariant derivative. The gluons obey Eq. 2.19, where here f^{abc} correspond to the structure constants

with $[\lambda^a, \lambda^b] = if^{abc}$. The strong coupling [27] $g_s = \sqrt{\alpha_s 4\pi}$ is a fundamental parameter of the SM. The running of α_s is described in Section 2.3.2.

At high momentum transfer (small distances), the strong force becomes weak, such that the binding between quarks e.g inside a proton can almost be neglected. At this stage, α_s is small enough to treat quarks as almost free particles. This observation is called *asymptotic freedom* [28]. With it, QCD calculations can then be factorized into a perturbative (hard process) and a non-perturbative process [29]. Quarks are never observed as free particles but rather only bounded, as hadrons. This is related to the fact, that at low energy scales (large distances) the coupling is so strong, that if two quarks move apart, it is at some point more energetically convenient to create new quark-antiquark pairs. This phenomenon, named *confinement*, is caused by the non-abelian structure of $SU(3)_c$, leading to self-interactions of gluons. In contrast to QED where the coupling decreases with increasing distance and photons do not carry electric charge, gluons carry color charge and can interact with themselves. The gluon self-interaction is shown in Fig. 2.3 for a three-gluon (left) and four-gluon (right) vertex.

Even though the formulation of QCD as a gauge theory was already born with the postulation of asymptotic freedom, the impact to take QCD predictions more seriously was then achieved and experimentally secured with the discovery of the gluons. Gluons were discovered at PETRA³ at DESY⁴ in 1979⁵ [30]. All detectors at this in these days world's biggest accelerator PETRA detected the process $e^+e^- \rightarrow q\bar{q}g$ [31]. The process is represented as Feynman diagram on the left in Fig. 2.4 and on the right a corresponding detected event by TASSO⁶ [32] with three hadron jets is shown.

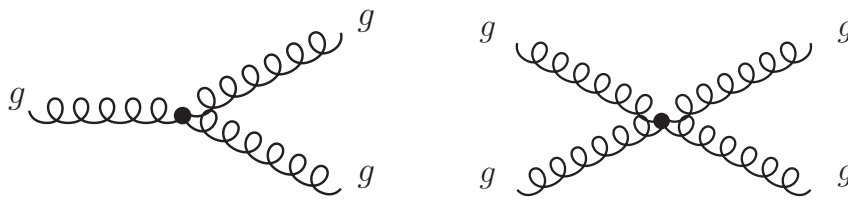


Fig. 2.3 Feynman diagrams for the gluon self-interaction. On the left (right) a three(four)-gluon vertex.

³Positron-Elektron-Tandem-Ring-Anlage

⁴The Deutsches Elektronen-Synchrotron is founded in 1959.

⁵These days physicists at DESY were not really aware of the discovery of the mediators of the strong force. In general, QCD was not taken that seriously in the field of particle physics in contrast to nowadays.

⁶TASSO is a collaboration of scientists who worked at DESY with the TASSO detector at PETRA The Prize of High Energy and Particle Physics was awarded to four members of that collaboration from CERN in 1995.

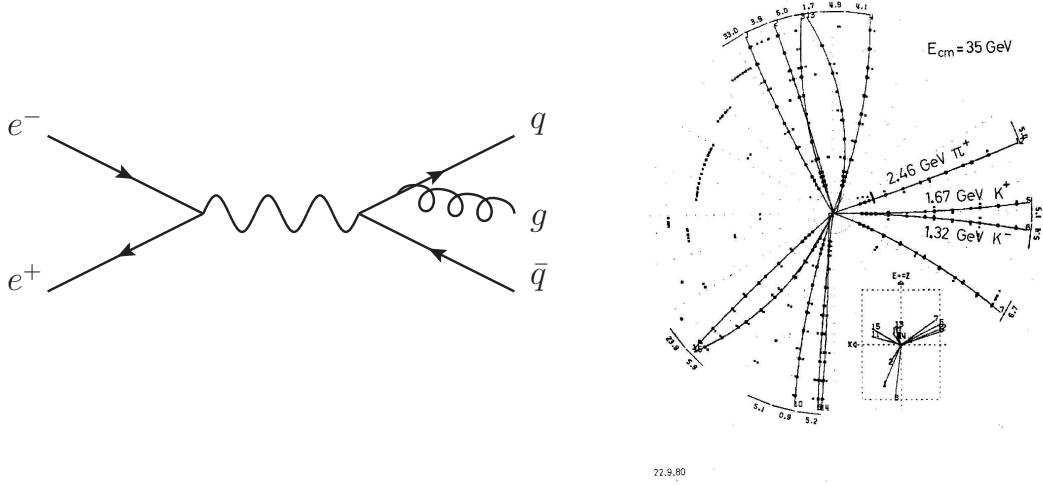


Fig. 2.4 On the left gluon radiation in a $e^+e^- \rightarrow q\bar{q}g$ process. On the right a detected collision of an electron and positron turning into two quarks which emit a gluon leading to 3 Jets. This event was detected by the TASSO detector at the accelerator PETRA at DESY and is taken from Reference [33].

2.2.5 The Higgs sector

Since this thesis deals with decays from massive fermions mediated by the massive bosons of the weak force, a brief description about the Higgs sector of the SM is given in the following. More descriptions can be found in Reference [34]. This section starts with showing how gauge bosons become massive in the SM and continues with the mass terms for fermions induced by the Higgs field and the Yukawa terms, respectively.

Massive gauge bosons

Gauge bosons become massive due to the spontaneous symmetry breaking caused by the Higgs potential. The Higgs Lagrangian in Eq. 2.2 is given by

$$\mathcal{L}_{Higgs} = (D^\mu \phi)^\dagger (D_\mu \phi) - V(\phi). \quad (2.20)$$

Here, D^μ corresponds to the covariant derivative given in Eq. 2.13. The Higgs potential $V(\phi)$ is given by

$$V(\phi) = -\mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2, \quad (2.21)$$

where μ and λ are real parameters. The generator of the $SU(2)_L$ group is given by $\tau^a = \sigma^a/2$ c.f. Eq. 2.13. The fields W_μ^a and B_μ are the gauge fields of the $SU(2)_L$ and $U(1)_Y$ gauge

group, respectively. A linear combination of these fields leads to the physical fields, c.f. Eq. 2.26. The Higgs weak isospin doublet is written as

$$\phi = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \quad (2.22)$$

For $\mu > 0$ and $\lambda > 0$ the Higgs potential develops a stable ground state with a non-zero vacuum expectation value, which is obtained by minimizing Eq. 2.21. The ground state is given by

$$\phi_0 = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}, \quad \text{with} \quad v = \sqrt{\frac{\mu^2}{\lambda}}. \quad (2.23)$$

After expanding the doublet Eq. 2.22 around the vacuum expectation value ϕ_0 and performing a gauge transformation one obtains

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \quad (2.24)$$

where $h(x)$ is the Higgs boson field and the remaining degree of freedom of the Higgs doublet. The mass terms of the gauge bosons $W^\pm$ and Z , arise from the kinetic term of the Higgs Lagrangian

$$\begin{aligned} (D^\mu \phi)^\dagger (D_\mu \phi) &= \frac{1}{2} (\partial_\mu h)^2 + \frac{1}{8} g^2 (v + h(x))^2 |W_\mu^1 + iW_\mu^2|^2 \\ &\quad + \frac{1}{8} (v + h(x))^2 |gW_\mu^3 - g'B_\mu|^2. \end{aligned} \quad (2.25)$$

After rewriting the fields W_μ^a and B_μ as

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (W_\mu^1 \mp iW_\mu^2), \quad Z_\mu = \frac{gW_\mu^3 - g'B_\mu}{\sqrt{g^2 + g'^2}}, \quad A_\mu = \frac{g'W_\mu^3 + gB_\mu}{\sqrt{g^2 + g'^2}}, \quad (2.26)$$

they turn into physical fields. The kinetic term Eq. 2.25 then becomes

$$(D^\mu \phi)^\dagger (D_\mu \phi) = \frac{1}{2} (\partial_\mu h)^2 + \left[\frac{1}{2} \left(\frac{g}{2} \right)^2 W^{\mu+} W_\mu^- + \frac{1}{2} \left(\frac{g^2 + g'^2}{4} \right) Z^\mu Z_\mu \right] (v + h)^2, \quad (2.27)$$

where the masses of $W_\mu^\pm$ and Z_μ appear in the bilinear terms of the fields and are given by

$$m_W = \frac{1}{2}vg \quad , \quad m_Z = \frac{1}{2}v\sqrt{g^2 + g'^2}. \quad (2.28)$$

Note that there is no mass term for the photon field A_μ , which corresponds to the fact that there is still an unbroken $U(1)_{\text{EM}}$ symmetry left. The Higgs boson mass is obtained by expanding the Higgs potential Eq. 2.21 around ϕ_0

$$\begin{aligned} V(\phi) &= -\mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \\ &= -\frac{\mu^2}{2} [v + h(x)]^2 + \frac{\lambda}{4} [v + h(x)]^4 \\ &= -\frac{\lambda v^4}{4} + \frac{1}{2} (2\lambda v^2) h(x)^2 + \lambda v h(x)^3 + \frac{\lambda}{4} h(x)^4 \\ &= -\frac{\lambda v^4}{4} + \underbrace{\frac{1}{2} m_h^2 h(x)^2}_{\text{mass term}} + \underbrace{\lambda v h(x)^3 + \frac{\lambda}{4} h(x)^4}_{\text{self couplings}}. \end{aligned} \quad (2.29)$$

The Higgs boson mass appears as a bilinear term in the Higgs field $h(x)$ and yields

$$m_h = \sqrt{2\lambda} v. \quad (2.30)$$

The expansion of the Higgs potential Eq. 2.29 also leads to double and triple Higgs self couplings, which are predictions of the theory. However, all fermions are still massless.

Fermion couplings to the Higgs boson

The mass terms of the massive fermions and their couplings to the Higgs boson are obtained within the electroweak symmetry breaking. The Yukawa term $\mathcal{L}_Y$ in Eq. 2.2 includes the interactions between massive fermions and the Higgs boson. In the following, the focus of $\mathcal{L}_Y$ lies on the quark sector, in particular on the charm- and beauty-quark as representatives for up- and down-type quarks. The Yukawa terms are written as

$$\begin{aligned} \mathcal{L}_{\text{Yukawa}}^c &= -Y_c \bar{q}_L \phi_c(x) c_R + \text{h.c.}, \\ \mathcal{L}_{\text{Yukawa}}^b &= -Y_b \bar{q}_L \phi(x) b_R + \text{h.c.}, \end{aligned} \quad (2.31)$$

where Y_q is the corresponding Yukawa coupling between a quark q and the Higgs boson and the chosen conjugated Higgs field

$$\phi_c(x) = i\sigma_2 \phi^\star(x), \quad (2.32)$$

and the left-handed SU(2) quark doublet of the third generation taken from Eq. 2.10

$$q_L = \begin{pmatrix} c \\ s \end{pmatrix}, \quad q'_L = \begin{pmatrix} t \\ b \end{pmatrix}. \quad (2.33)$$

In Eq. 2.31 c_R and b_R correspond to the right-handed SU(2) quark singlet c.f. Eq. 2.11 and $Y_{c,b}$ to the individual Yukawa couplings of charm- and beauty quarks. After expanding the Higgs field around the vacuum expectation value ϕ_0 one obtains the mass terms of the fermions and their couplings to the Higgs boson. The quarks contain the same Higgs field ϕ_c , which is charge conjugated. Thus, the Yukawa couplings for up- and down-type quarks both scale with the same vacuum expectation value v . Inserting Eq. 2.24 into the Yukawa term Eq. 2.31 for the up-type quark, one obtains

$$\begin{aligned} \mathcal{L}_{\text{Yukawa}}^c &= -Y_c \frac{1}{\sqrt{2}} (\bar{c}_L, \bar{b}_L) \begin{pmatrix} v + h(x) \\ 0 \end{pmatrix} c_R + \text{h.c.} \\ &= -Y_c \frac{1}{\sqrt{2}} [v + h(x)] \bar{c}_L c_R + \text{h.c.} \\ &= \underbrace{-m_c \bar{c}_L c_R}_{\text{mass term}} - \underbrace{\frac{m_c}{v} \bar{c}_L c_R h(x)}_{\text{coupling term}} + \text{h.c.}, \end{aligned} \quad (2.34)$$

with the charm-quark mass

$$m_c = \frac{1}{\sqrt{2}} Y_c v. \quad (2.35)$$

The calculation for the beauty quark is analogous and leads to the beauty-quark mass

$$m_b = \frac{1}{\sqrt{2}} Y_b v. \quad (2.36)$$

The masses for charm and beauty quarks m_q are part of the QCD Lagrangian c.f. Eq. 2.17 and enter the cross section measurement in hadron hadron collisions studied in this thesis.

2.3 QCD Cross Section in Hadron Hadron Collisions

QCD predictions can be tested by scattering experiments for instance with a hadron collider like the LHC. Within this thesis measurements of certain QCD cross sections are obtained. How these predictions can be calculated will be explained in this section.

Starting with the QCD Lagrangian Eq. 2.17 a set of partial differential equations can be derived which can be calculated in perturbation theory. In this case, a perturbative cross section for instance for a scattering process can be written as Taylor expansion in the strong coupling α_s

$$\hat{\sigma} = \sigma_0 \alpha_s^0 + \sigma_1 \alpha_s^1 + \sigma_2 \alpha_s^2 + \dots = \sum_{i=0}^n \sigma_i \alpha_s^i, \quad (2.37)$$

where n represents the order of the calculation. If n is the smallest integer where the coefficient σ_i is nonzero, the calculation is performed at *leading order* (LO), while an increase of the order leads to *next-to-leading-order* (NLO) and *next-to-next-to-leading-order* (NNLO), et cetera.

In general, a measurable cross section in particle collisions is given by

$$\sigma = \frac{N}{L_{int}}, \quad (2.38)$$

where σ corresponds to the cross section, N to the number of counted collisions of a certain process and L_{int} to the integrated luminosity. There are different types of cross sections. If in addition to the considered final state other collision products are included in the calculation, it is called *inclusive* cross section, and *exclusive* otherwise. Cross sections can be measured *differentially* in kinematic variables like the transverse momentum but they can be also measured as a function of geometric variables like the detection angle. However, a purely perturbative calculation is often not possible and thus, phenomenological approaches are needed. In this case, the QCD calculation contains a non-perturbative component.

In perturbation theory cross sections of a QCD process in proton-proton collisions can then often be factorized. By considering the process $P_1 P_2 \rightarrow \mathcal{H} X$ where P_1, P_2 are protons with the momentum $\vec{p}_1, \vec{p}_2$, $\mathcal{H}$ a hadron and X some remnant, the cross section can be written as

$$\sigma(P_1 P_2 \rightarrow \mathcal{H} X)_{\text{QCD}} = \int d^3 p dx_n dx_l dz_m \sum_{n,l,m} f_n(x_n) f_l(x_l) |\mathcal{M}(nl \rightarrow mX)|^2 F_m^{\mathcal{H}}(z_m), \quad (2.39)$$

$$= \sum_{nl} \int dx_n dx_l \underbrace{f_n(x_n, \mu_F^2) f_l(x_l, \mu_F^2)}_{\text{PDFs}} \cdot \underbrace{\hat{\sigma}(x_n \vec{p}_1, x_l \vec{p}_2, \mu_F^2, \mu_R^2)}_{\text{hard process}} \otimes \text{hadronization} \quad (2.40)$$

where n, l and m stand for the partons (quarks or gluons) involved in that process. The matrix element $\mathcal{M}(nl \rightarrow mX)$ corresponds to the amplitude for the partonic process $nl \rightarrow mX$. The parton distribution functions (PDF) are $f_n(x_n), f_l(x_l)$ representing the probability to find a parton n, l in the proton $p_{1,2}$ with a fraction x_n, x_l of the proton momentum $\vec{p}_1, \vec{p}_2$. The parton fragmentation function is given by $F_m^{\mathcal{H}}(z_m)$. It corresponds to the probability that a parton m hadronises into $\mathcal{H}$ carrying the momentum fraction z_m of the parton m . The parameters μ_F and μ_R in Eq. 2.40 are the factorization scale and the renormalization scale, respectively, explained in Section 2.3.2. The product of the PDFs and the hard process is convoluted with the hadronization. An illustration of Eq. 2.40 is shown in 2.5. In the following subsections, the three parts of that equation are explained in more detail.

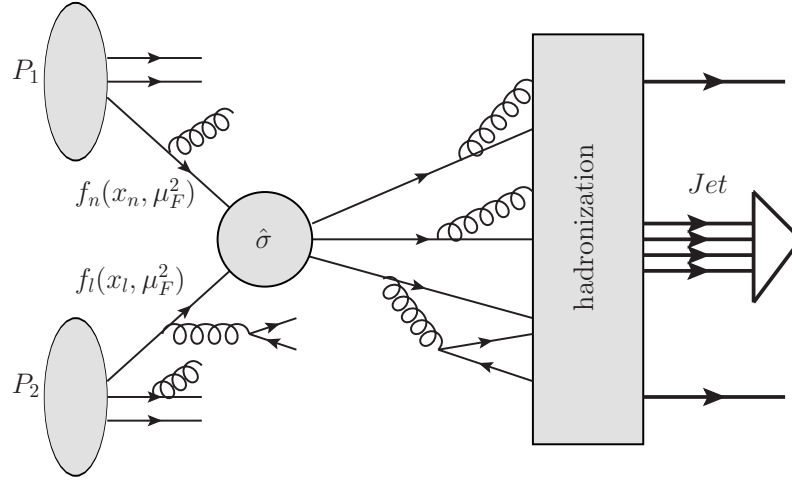


Fig. 2.5 Illustration of the cross section calculation in a proton proton collision. It can be split into a product of PDFs and the hard process convoluted with the hadronization. The PDFs are $f_n(x_n, \mu_F)$ and $f_l(x_l, \mu_F)$ where $x_{i=n,l}$ correspond to the momentum fraction taken from the proton (P_1, P_2) and μ_F and μ_R are the factorization scale and the renormalization scale, respectively. The hard process is represented by $\hat{\sigma}$ and the box represents the hadronization.

2.3.1 Parton model and PDFs

Measurements of differential cross sections lead to the observation, that protons are not point-like particles but consist of gluons and quarks, which are called *partons*. This *parton model* says that a proton consists of three valence quarks, quark-anti-quark pairs (sea quarks) and gluons. The Drell-Yan process $pp \rightarrow \ell^+ \ell^- + X$ can be explained by the parton model and is shown in Fig. 2.6. Here, a quark annihilates with an anti-quark to a virtual photon or Z boson, which decays into a pair of leptons. The remnant of the protons is presented with X and X' , in Fig. 2.6. The probability that a parton for instance a sea quark takes part in that

process is described by the parton distribution functions (PDFs). There are different PDFs for the three types of partons (valence quarks, sea quarks and gluons). Turning on QCD, the parton densities become scale dependent, and are to be evaluated at the scale μ_F in Eq. 2.40. The universal properties of the proton can be described by these PDFs and can be measured with different experiments, while some of these experiments have different sensitivity to the PDFs for gluons, valence quarks or sea quarks [35].

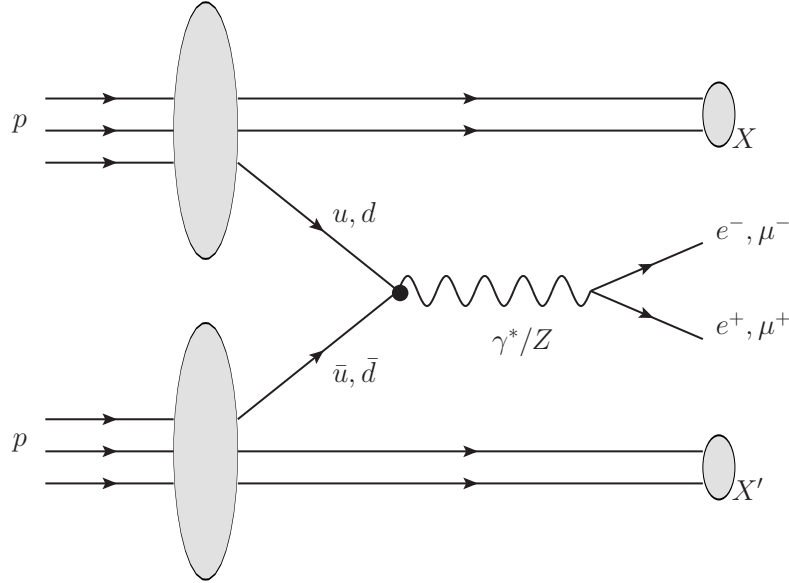


Fig. 2.6 Collision of two protons at high momentum transfer. The Drell-Yan process $pp \rightarrow \ell^+ \ell^- + X$ is a process used to study hadron hadron collisions and can be theoretically calculated since quarks can be treated as free particles at high energy scales. It is possible to factorize the calculation into a hard and soft process.

For instance, the Drell-Yan data is used in PDF fits since this process is sensitive to the distribution of anti-quarks. Next to Drell-Yan, fixed target data or deep inelastic scattering (DIS⁷) processes like HERA⁸ data are mostly included for the determination of a certain PDF. To achieve best results, a global fit with data from different measurements is done. Divergences in the perturbative parton splitting are handled by the non physical parameter μ_F described in Section 2.3.2. The results of a fixed order calculation depend on the chosen value of μ_F , and this dependence is reduced⁹ with increasing order of α_s c.f. Eq. 2.37. To

⁷Deep inelastic scattering is used to describe scattering processes that probe the structure of a hadron by colliding a lepton with it.

⁸Hadron-Electron Ring Accelerator at DESY.

⁹This means when this full series is calculated for a process, the dependency would fully vanish reflecting the fact, that those parameters are non-physical and the observable should not depend on them.

evolve PDFs from one scale μ_F to a different scale, the DGLAP¹⁰ [36–38] equations are used. Nowadays there are several groups with different approaches to investigating parton distributions. Existing PDF fits are for instance HERAPDF 1.0 [39], HERAPDF 2.0 [40], CTEQ6.6 [41], NNPDF2.0 [42] or NNPDF3.1 [43]. For the latter one an example for the different PDFs at NNLO is shown in Fig. 2.7. More about PDFs can be found for instance in Reference [44].

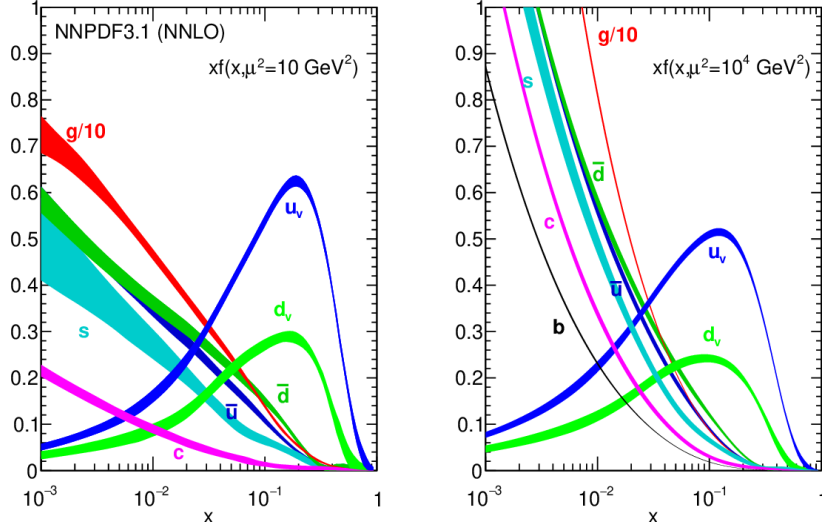


Fig. 2.7 The NNLO PDF for different values of μ_F from the NNPDF3.1 analysis. These set of PDFs were evaluated for (left) $\mu_F^2 = 10 \text{ GeV}^2$ and (right) $\mu_F^2 = 10^4 \text{ GeV}^2$. For lower momentum fractions x the contribution from sea quarks or gluons is dominating and for higher values of x for valence quarks. Taken from Reference [43].

2.3.2 Hard process and factorization

The hard processes $\hat{\sigma}$ at high scales lead typically to a large momentum transfer or large invariant mass of particles in the final state. The probability for a hard process to occur is proportional to the square of the transition amplitude $|\mathcal{M}|^2$ for a partonic process (c.f. Eq. 2.39), and can be calculated in perturbation theory [45]. The transition amplitude, which can be represented in terms of Feynman diagrams, can be calculated from the QCD Lagrangian given in Eq. 2.17. The calculation of $\mathcal{M}$ and $\hat{\sigma}$, respectively, is in general done at a fixed order of Eq. 2.37 for instance at LO or NLO.

¹⁰Evolution equations for QCD splittings are also known as DGLAP equations which were first published by Altarelli–Parisi in 1977.

To a fixed order calculation a lot of Feynman diagrams can contribute, of which some contain loops or radiation of massless particles as shown in Fig. 2.2 and Fig. 2.4. Even if individual contributions can be divergent, the sum of all diagrams has to be finite. It can be distinguished between ultraviolet (UV) and infrared (IR) divergences.

UV divergences are caused by very large momenta inside a loop of a Feynman diagram contributing to the transition amplitude. In a fixed order calculation these divergences are absorbed into experimentally measurable quantities. This procedure is called renormalization. The introduction of the energy scale parameter μ_R , which handles the UV divergences, is separating the infinite part from the finite part. This leads to α_s becoming a function of μ_R , which is called *running* of the strong coupling. μ_R is then the scale at which α_s will be evaluated in the hard matrix element.

IR divergences are caused by the fact that a virtual or real particle can carry a vanishing momentum or be emitted at a vanishing angle. This kind of divergences cancels at any order of the expanded cross section (c.f. Eq. 2.37) between the different contributions.¹¹ The introduction of μ_F treats IR divergences in the initial state, separates the low from the high energy process and leads to the PDFs and fragmentation functions becoming a function of μ_F .

These energy scales are not physical parameters and an observable like the cross section should not depend on them. By calculating the total sum of the perturbative expansion c.f. Eq. 2.37 this would be the case, but for a fixed order the dependency decreases only by computing more and more terms. The higher the order the less the observables depend on these introduced scales. A monotonous dependency for observables is found at LO, but even at NLO the dependency does not fully vanish. The value of μ_F and μ_R can for instance be set to $\mu_R = \mu_F = \mu_0 = \sqrt{m^2 + p_T^2}$, where m and p_T are the mass and transverse momentum from the parton (quark). Another way is to find a value for these scales such that the underlying observables do not vary much by a variation of these scales. Since they are not physical parameters and observables should theoretically not depend on them, any ansatz is generally feasible. It is a convention that a theoretical uncertainty is estimated by a variation of μ_R and μ_F individually by 2 and $\frac{1}{2}$ [46, 47] leading to seven variations as illustrated in Fig. 2.8.

A pure measurement of $\hat{\sigma}$ is not possible, instead always of σ_{QCD} c.f. Eq. 2.40. The full prediction of σ_{QCD} in perturbation theory is complex due to the fact that the energy scale becomes lower and lower due to the radiation of particles like gluons or gluon-splitting into

¹¹Kinoshita-Lee-Nauenberg theorem: infrared divergences caused by loop integrals cancel with the ones coming from phase space integral.

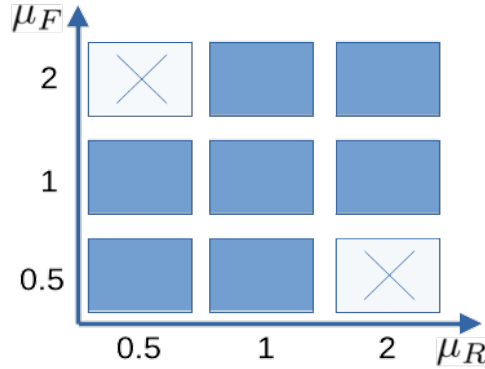


Fig. 2.8 In blue seven combinations of the variation of the renormalization and factorization scale μ_R and μ_F , respectively, are shown. The combinations are marked which are not considered in the scale uncertainty estimation.

quarks or gluons, which in turn can lead to many particles in the final state. The probability for gluon-splitting or of gluon radiation diverges in the soft and collinear limit [48]. By collinear it is meant that a gluon splits into two partons at a low angle and soft means that gluons are emitted with low transverse momentum.

If the energy scale is low enough, partons can not be treated anymore as free particles such that perturbation theory is not possible anymore. These non perturbative effects have to be adjusted to data and models discussed in the following section are needed.

2.3.3 Parton shower and hadronization

An advantage of the factorization is that QCD splittings, which lead to divergences in the soft and collinear limit, can be calculated independently from the calculation of the hard process. Apart from the PDFs, this kind of splittings can be treated by a procedure called *parton shower*. If the calculation of the hard process is performed for a certain scale, a parton shower can be added. There are different approximations of showers that differ in their ordering and handling, there are for instance angular and momentum ordered showers [49, 50] and so called dipole showers include color connections between showering partons [51].

All parton showers have in common the implementation of the Sudakov form factor¹², which describes the non-emission probability of partons for a certain scale of μ_F in an algorithm describing the parton shower. It works recursively by decreasing the value of μ_F per iteration. In this way, particles can be added to the hard process via a matching algorithm [52], which prevents double counting by the combining procedure of the perturbative calculation

¹²The Sudakov form factor describes the emission frequency in the parton evolution and it is derived from a differential equation, which contains the DGLAP equations.

with the non-perturbative part. This matching procedure is explained in Section 4.7 and for NLO QCD calculation implemented in Monte Carlo (MC) codes. These procedures are implemented in common NLO MC codes used in HEP¹³ calculations [53] such as POWHEG [54–56] or HERWIG [51, 57].

At one point the energy scale becomes so low that all partons are not treated anymore as free particles. Then, they start to form hadrons (mesons or baryons) in a process called hadronization. This procedure needs to be modeled and is simulated for example with PYTHIA [58, 59](LO+PS) or HERWIG(NLO+PS). In the end, the final state consists only of stable particles, which means that those hadrons which are not stable are decayed. These models have a lot of free parameters which need to be tuned by comparing them with experimental measurements.

¹³High Energy Physics

References

- [1] A. Salam and J. C. Ward, “On a gauge theory of elementary interactions”, *Il Nuovo Cimento (1955-1965)* **19** (Jan, 1961) 165–170, 10.1007/BF02812723.
- [2] S. L. Glashow, “Partial-symmetries of weak interactions”, *Nuclear Physics* **22** (1961), no. 4, 579–588, [https://doi.org/10.1016/0029-5582\(61\)90469-2](https://doi.org/10.1016/0029-5582(61)90469-2).
- [3] S. Weinberg, “A Model of Leptons”, *Phys. Rev. Lett.* **19** (Nov, 1967) 1264–1266, 10.1103/PhysRevLett.19.1264.
- [4] ATLAS Collaboration, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”, *Phys. Lett.* **B716** (2012) 1–29, 10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
- [5] CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, *Phys. Lett.* **B716** (2012) 30–61, 10.1016/j.physletb.2012.08.021, arXiv:1207.7235.
- [6] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons”, *Phys. Rev. Lett.* **13** (Aug, 1964) 321–323, 10.1103/PhysRevLett.13.321.
- [7] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons”, *Phys. Rev. Lett.* **13** (Oct, 1964) 508–509, 10.1103/PhysRevLett.13.508.
- [8] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global Conservation Laws and Massless Particles”, *Phys. Rev. Lett.* **13** (Nov, 1964) 585–587, 10.1103/PhysRevLett.13.585.
- [9] M. Czakon, P. Fiedler, and A. Mitov, “Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$ ”, *Phys. Rev. Lett.* **110** (2013) 252004, 10.1103/PhysRevLett.110.252004, arXiv:1303.6254.
- [10] LHCTopWG, “LHCTopWG”, <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWG> (Urldate:23.08.2021) <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWG>.
- [11] D. J. Griffiths, “Introduction to elementary particles; 2nd rev. version”. Physics textbook. Wiley, New York, NY, 2008.
- [12] P. D. Group, “Review of Particle Physics”, *Progress of Theoretical and Experimental Physics* **2020** (08, 2020) 10.1093/ptep/ptaa104. 083C01.
- [13] F. Mandl and G. Shaw, “Quantum field theory”. Wiley, Chichester, 2. ed. edition, 2010.
- [14] J. C. Romao and J. P. Silva, “A resource for signs and Feynman diagrams of the Standard Model”, *Int. J. Mod. Phys.* **A27** (2012) 1230025, 10.1142/S0217751X12300256, arXiv:1209.6213.
- [15] C. Berger, “Elementarteilchenphysik : Von den Grundlagen zu den modernen Experimenten”. Springer-Lehrbuch. Springer, Berlin, 3., aktualisierte und überarb. aufl. edition, 2014.

- [16] I. Falconer, “J J Thomson and the discovery of the electron”, *Physics Education* **32** (jul, 1997) 226–231, 10.1088/0031-9120/32/4/015.
- [17] “Standard Model”. https://en.wikipedia.org/wiki/Standard_Model.
- [18] R. P. Feynman, “The Theory of Positrons”, *Phys. Rev.* **76** (Sep, 1949) 749–759, 10.1103/PhysRev.76.749.
- [19] M. E. Peskin and D. V. Schroeder, “An introduction to quantum field theory”. Westview, Boulder, CO, 1995. Includes exercises.
- [20] D. Binosi and L. Theussl, “JaxoDraw: A Graphical user interface for drawing Feynman diagrams”, *Comput. Phys. Commun.* **161** (2004) 76–86, 10.1016/j.cpc.2004.05.001, arXiv:hep-ph/0309015.
- [21] A. Salam, “Weak and Electromagnetic Interactions”, in *Elementary particle theory*, N. Svartholm, ed., p. 367–377. Almquist & Wiksell, 1968.
- [22] A. Djouadi, “The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model”, *Phys. Rept.* **457** (2008) 1–216, 10.1016/j.physrep.2007.10.004, arXiv:hep-ph/0503172.
- [23] T.-P. Cheng and L.-F. Li, “Gauge theory of elementary particle physics”. Oxford science publications. Clarendon Press, Oxford [u.a.], repr. edition, 2008.
- [24] A. Salam, “Weak and electromagnetic interactions”, pp. 244–254. arXiv:https://www.worldscientific.com/doi/pdf/10.1142/9789812795915_0034. 10.1142/9789812795915_0034.
- [25] M. Gell-Mann, “The interpretation of the new particles as displaced charge multiplets”, *Il Nuovo Cimento (1955-1965)* **4** (1956), no. 2, 848–866, 10.1007/BF02748000.
- [26] T. Nakano and K. Nishijima, “Charge Independence for V-particles*”, *Progress of Theoretical Physics* **10** (11, 1953) 581–582, 10.1143/PTP.10.581, arXiv:<https://academic.oup.com/ptp/article-pdf/10/5/581/5364926/10-5-581.pdf>.
- [27] G. Dissertori, “The Determination of the Strong Coupling Constant”, *Adv. Ser. Direct. High Energy Phys.* **26** (2016) 113–128, 10.1142/9789814733519_0006, arXiv:1506.05407.
- [28] G. ’t Hooft, “When was asymptotic freedom discovered? or The rehabilitation of quantum field theory”, *Nuclear Physics B - Proceedings Supplements* **74** (1999), no. 1, 413–425, [https://doi.org/10.1016/S0920-5632\(99\)00207-8](https://doi.org/10.1016/S0920-5632(99)00207-8). QCD 98.
- [29] J. C. Collins, D. E. Soper, and G. F. Sterman, “Factorization of Hard Processes in QCD”, *Adv. Ser. Direct. High Energy Phys.* **5** (1989) 1–91, 10.1142/9789814503266_0001, arXiv:hep-ph/0409313.
- [30] R. Brandelik et al., “Evidence for planar events in e+e- annihilation at high energies”, *Physics Letters B* **86** (1979), no. 2, 243–249, [https://doi.org/10.1016/0370-2693\(79\)90830-X](https://doi.org/10.1016/0370-2693(79)90830-X).

- [31] D. P. Barber et al., “Discovery of Three-Jet Events and a Test of Quantum Chromodynamics at PETRA”, 10.1103/PhysRevLett.43.830.
- [32] R. Brandelik et al., “Evidence for a spin-1 gluon in three-jet events”, *Physics Letters B* **97** (December, 1980) 453–458, 10.1016/0370-2693(80)90639-5.
- [33] DESY, “40 years of gluon”, Urldate:10.08.2021.
https://www.desy.de/news/backgrounders/40_years_of_gluon.
- [34] R. Wolf, “The Higgs Boson Discovery at the Large Hadron Collider”. Springer Tracts in Modern Physics ; 264SpringerLink : Bücher. Springer, Cham, 2015.
- [35] D. E. Soper, “Parton distribution functions”, *Nucl. Phys. B Proc. Suppl.* **53** (1997) 69–80, 10.1016/S0920-5632(96)00600-7, arXiv:hep-lat/9609018.
- [36] Y. L. Dokshitzer, “Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics.”, *Sov. Phys. JETP* **46** (1977) 641–653.
- [37] V. N. Gribov and L. N. Lipatov, “Deep inelastic $e p$ scattering in perturbation theory”, *Sov. J. Nucl. Phys.* **15** (1972) 438–450.
- [38] G. Altarelli and G. Parisi, “Asymptotic freedom in parton language”, *Nuclear Physics B* **126** (1977), no. 2, 298–318, [https://doi.org/10.1016/0550-3213\(77\)90384-4](https://doi.org/10.1016/0550-3213(77)90384-4).
- [39] H1, ZEUS Collaboration, “Combined Measurement and QCD Analysis of the Inclusive e^+p Scattering Cross Sections at HERA”, *JHEP* **01** (2010) 109, 10.1007/JHEP01(2010)109, arXiv:0911.0884.
- [40] H1, ZEUS Collaboration, “Combination of measurements of inclusive deep inelastic $e^\pm p$ scattering cross sections and QCD analysis of HERA data”, *Eur. Phys. J. C* **75** (2015), no. 12, 580, 10.1140/epjc/s10052-015-3710-4, arXiv:1506.06042.
- [41] P. M. Nadolsky et al., “Implications of CTEQ global analysis for collider observables”, *Phys. Rev. D* **78** (2008) 013004, 10.1103/PhysRevD.78.013004, arXiv:0802.0007.
- [42] R. D. Ball et al., “A first unbiased global NLO determination of parton distributions and their uncertainties”, *Nucl. Phys. B* **838** (2010) 136–206, 10.1016/j.nuclphysb.2010.05.008, arXiv:1002.4407.
- [43] NNPDF Collaboration, “Parton distributions from high-precision collider data”, *Eur. Phys. J. C* **77** (2017), no. 10, 663, 10.1140/epjc/s10052-017-5199-5, arXiv:1706.00428.
- [44] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, “Parton distributions for the LHC”, *Eur. Phys. J. C* **63** (2009) 189–285, 10.1140/epjc/s10052-009-1072-5, arXiv:0901.0002.
- [45] G. Sterman et al., “Handbook of perturbative QCD”, *Rev. Mod. Phys.* **67** (Jan, 1995) 157–248, 10.1103/RevModPhys.67.157.

- [46] M. Cacciari, M. Greco, and P. Nason, “The P(T) spectrum in heavy flavor hadroproduction”, *JHEP* **05** (1998) 007, 10.1088/1126-6708/1998/05/007, arXiv:hep-ph/9803400.
- [47] M. Cacciari, S. Frixione, and P. Nason, “The p(T) spectrum in heavy flavor photoproduction”, *JHEP* **03** (2001) 006, 10.1088/1126-6708/2001/03/006, arXiv:hep-ph/0102134.
- [48] G. P. Salam, “Elements of QCD for hadron colliders”, in *2009 European School of High-Energy Physics*. 11, 2010. arXiv:1011.5131.
- [49] A. Banfi, G. Corcella, and M. Dasgupta, “Angular ordering and parton showers for non-global QCD observables”, *JHEP* **03** (2007) 050, 10.1088/1126-6708/2007/03/050, arXiv:hep-ph/0612282.
- [50] Z. Nagy and D. E. Soper, “Ordering variable for parton showers”, *JHEP* **06** (2014) 178, 10.1007/JHEP06(2014)178, arXiv:1401.6366.
- [51] G. Corcella et al., “HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)”, *JHEP* **01** (2001) 010, 10.1088/1126-6708/2001/01/010, arXiv:hep-ph/0011363.
- [52] J. Alwall et al., “Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions”, *Eur. Phys. J. C* **53** (2008) 473–500, 10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.
- [53] S. Höche, “Introduction to parton-shower event generators”, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders*, pp. 235–295. 2015. arXiv:1411.4085. 10.1142/9789814678766_0005.
- [54] E. Bagnaschi, G. Degrandi, P. Slavich, and A. Vicini, “Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM”, *JHEP* **02** (2012) 088, 10.1007/JHEP02(2012)088, arXiv:1111.2854.
- [55] S. Alioli, P. Nason, C. Oleari, and E. Re, “NLO Higgs boson production via gluon fusion matched with shower in POWHEG”, *JHEP* **04** (2009) 002, 10.1088/1126-6708/2009/04/002, arXiv:0812.0578.
- [56] C. Oleari, “The POWHEG-BOX”, *Nucl. Phys. Proc. Suppl.* **205-206** (2010) 36–41, 10.1016/j.nuclphysbps.2010.08.016, arXiv:1007.3893.
- [57] M. Bahr et al., “Herwig++ Physics and Manual”, *Eur. Phys. J. C* **58** (2008) 639–707, 10.1140/epjc/s10052-008-0798-9, arXiv:0803.0883.
- [58] T. Sjöstrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual”, *JHEP* **05** (2006) 026, 10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.
- [59] T. Sjöstrand, S. Mrenna, and P. Skands, “A brief introduction to PYTHIA 8.1”, *Computer Physics Communications* **178** (2008), no. 11, 852 – 867, <https://doi.org/10.1016/j.cpc.2008.01.036>.

Chapter 3

Physics in CMS with the LHC at CERN

Contents

3.1	CERN	32
3.1.1	The LHC	32
3.2	The Compact Muon Solenoid Detector	33
3.2.1	The coordinate system of CMS	36
3.2.2	Magnet	36
3.2.3	Muon system	37
3.2.4	Hadronic calorimeter	38
3.2.5	Electromagnetic calorimeter	38
3.3	The CMS Tracker and its Alignment	39
3.3.1	The silicon tracker	39
3.3.2	Tracker alignment	41
	References	47

This chapter provides a condensed description of the LHC (Large Hadron Collider), its different experiments based at CERN¹ and in particular of the Compact Muon Solenoid (CMS), with which the data analysed in this thesis were collected.

¹Originally ‘Conseil Européen pour la Recherche Nucléaire’, now ‘Organisation Européenne pour la Recherche Nucléaire’

3.1 CERN

Since 1954 the European Organization for Nuclear Research (CERN) is a hot spot for particle physicists. In the field of particle physics, several discoveries were made like the discovery of the massive electroweak bosons $W^\pm$ and Z^0 with the SPS (Super Proton Synchrotron) [1] shown in Fig. 3.1. By today several developments were made at CERN with impacts beyond high energy physics. The world wide web [2] is one of the most prominent inventions developed for faster interaction among CERN scientists around the globe.

CERN, based in Geneva, is the largest particle physics laboratory in the world and the organisation consists of 23 member states. The current detectors were designed to measure high energetic colliding particles at the LHC, which are accelerated through different accelerators shown in Fig. 3.1. By today various experiments recorded hundreds of petabytes of data. In summer of 2012 one of the most notable achievements in particle physics was made: the Higgs boson discovery [3, 4], which was one of the reasons for building the LHC.

3.1.1 The LHC

The LHC was built between 1998 and 2008 in the former LEP² tunnel to accelerate hadrons, has a circular circumference of 26.7 km and is located at a depth of 50 m to 175 m underground. It is originally designed to accelerate protons to center of mass energies up to $\sqrt{s} = 14$ TeV and heavy ion collisions up to $\sqrt{s} = 5.5$ TeV. In the following, a description of the latter is omitted, since the study of this work deals exclusively with pp collisions.

Remarkably, a small container filled with hydrogen gas serves as the proton source for the LHC. Before protons are accelerated with the LHC, they pass through several pre-accelerators, which is shown schematically in Fig. 3.1. First, they are accelerated to an energy of $\approx$ tens of MeV using a linear accelerator called Linac2 and then using the Proton Synchrotron Booster (PBS) together with the Proton Synchrotron (PS) to accelerate them to the energy of 25 GeV. The last accelerator before the LHC is the SPS, with which protons are then brought to the energy of 450 GeV and then injected into the LHC via two tunnels. At the LHC there are four interaction points where the proton beams collide and the associated detectors are located. There are four main experiments at the LHC, shown in Fig. 3.1 with yellow dots, with their detectors named:

- **CMS** (Compact Muon Solenoid) is a compact multi-purpose detector designed to ensure a high muon resolution [5].

²Large Electron Positron

- **ATLAS** (A Toroidal LHC ApparatuS) is a large multi-purpose detector [6].
- **ALICE** (A Large Ion Collider Experiment) is a detector designed to study heavy ion collisions and in particular quark-gluon plasma production [7].
- **LHCb** (Large Hadron Collider beauty) is a fully forward detector designed to study beauty physics [8].

Accelerators are not only characterized by their shape or length, but also by the instantaneous luminosity $\mathcal{L}$ indicating the ratio of the number of expected events to the associated cross section for the underlying process. It is given by

$$\sigma = \frac{dN}{dt} \frac{1}{\mathcal{L}}, \quad (3.1)$$

with the fully accelerator dependent parameter

$$\mathcal{L} = \frac{n_b N_1 N_1 f}{4\pi\sigma_x\sigma_y}. \quad (3.2)$$

Here, n_b corresponds to the number of proton bunches per beam with N_1 and N_1 the number of protons in each bunch. The frequency with which these bunches are colliding is given by f and the transverse dimension of each bunch by σ_x and σ_y . The integral of the luminosity quantifies the amount of collision data corresponding to the integral of the instant luminosity such that Eq. 3.1 can be written as

$$\sigma = \frac{N}{L_{int}}, \quad L_{int} = \int dt \mathcal{L}. \quad (3.3)$$

In this thesis data from the LHC Run 2 in the year 2015 at a center of mass energy of 5.02 TeV with an integrated luminosity of $L_{int} = 40.17 \text{ nb}^{-1}$ is analysed. A more precise description of this data and its use can be found in Section 6.1. More information about luminosity measurements within CMS can be found in References [9, 10].

3.2 The Compact Muon Solenoid Detector

The multi-purpose CMS detector with its diameter of 14.6m, length of 21.6m and weight of 14000 tons is a very heavy and compact one, compared to the ATLAS detector with 46m length, a radius of 12.5m and a weight of 7000 tons. Protons which are accelerated by the

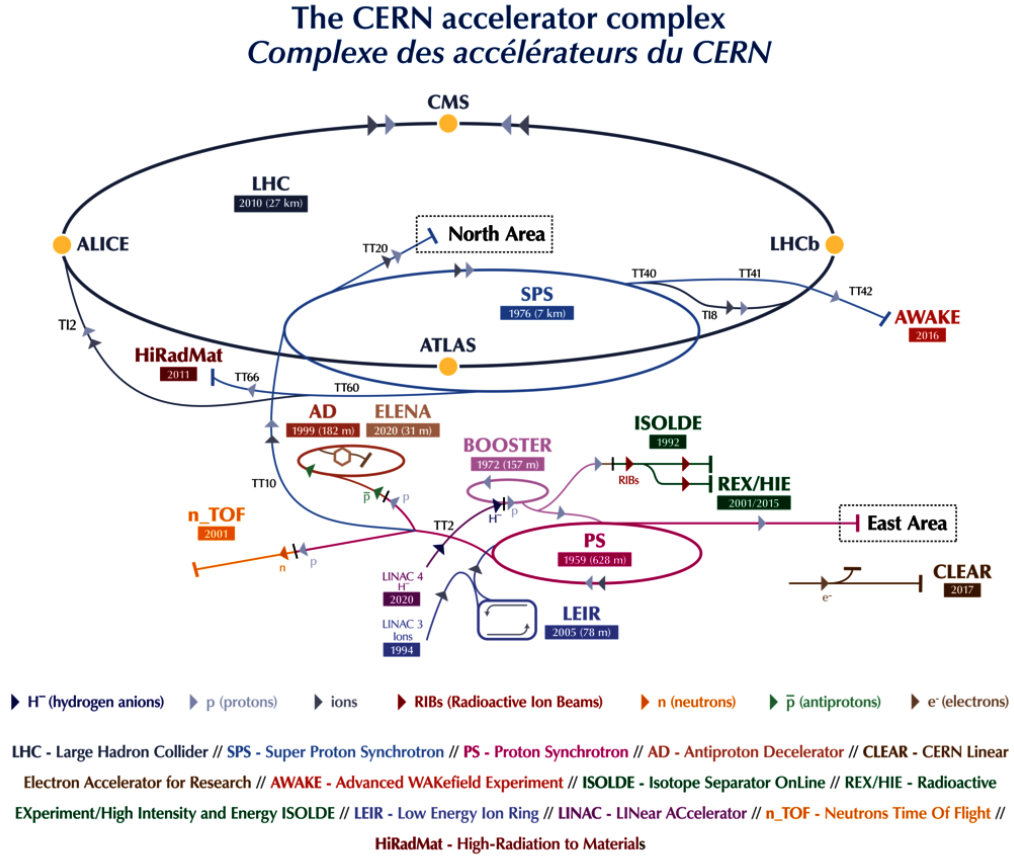


Fig. 3.1 A schematic view of the LHC and its pre-accelerators at CERN. Before protons are injected into the LHC (deep blue), they traverse different pre-accelerators. First, protons are accelerated with the LINAC (purple), followed by the Booster (light purple), the Proton Synchrotron (pink), and the Super Proton Synchrotron (light blue). The four main experiments at the LHC (CMS, ATLAS, ALICE, LHCb) are represented with yellow dots. This figure is taken from Reference [11].

LHC are colliding in the center of the detector. The decay products scatter in all directions, while different detector layers react differently to the particle types.

Starting from the outside, the detector consists of the muon chambers, the hadronic calorimeter (HCAL), the electromagnetic calorimeter (ECAL) and the silicon tracker. The presence of the magnetic field induced by the superconducting solenoid is manifested by the curvature of the muon track shown in Fig. 3.2 representing a slice of the CMS detector. Here, a charged particle (muon) is changing its direction due to the solenoid with a strong magnetic field of 3.8 T while passing transversely through each detector layer.

In the following each sub-detector is described and since this analysis relies mostly on the track reconstruction, the discussion of each sub-detector (layer) is starting from the outer layer and ends in the innermost CMS detector, the tracker. The following discussion will be kept brief, while a more precise description can be found in References [12–14].

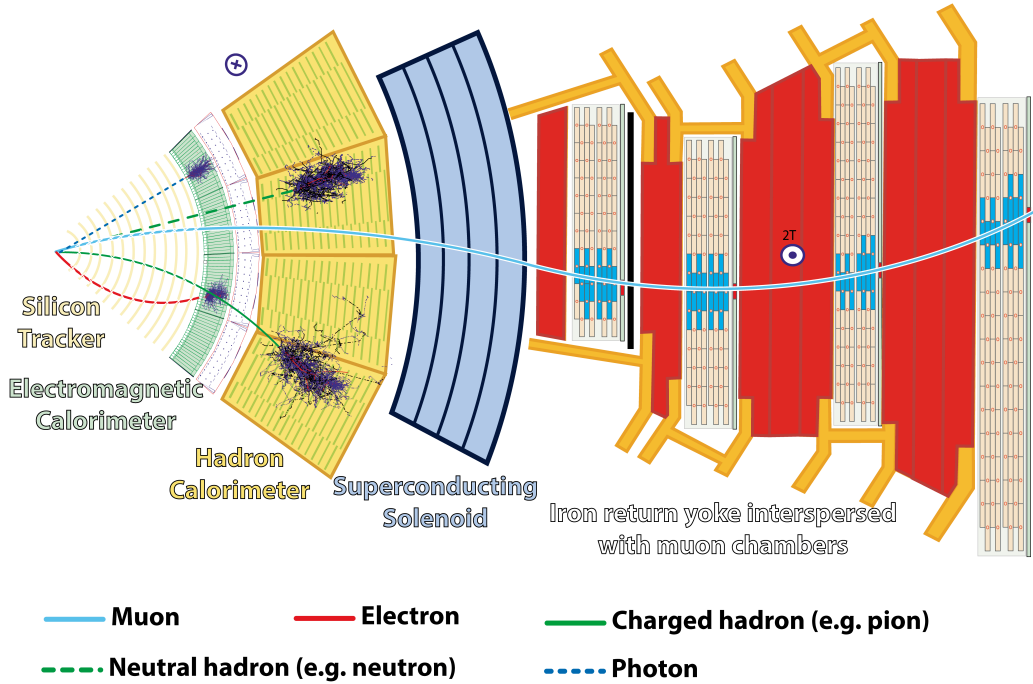


Fig. 3.2 A schematic view of the CMS and its sub-detectors. Starting from the outer layers it shows the muon chambers, the hadronic calorimeter (HCAL), the electromagnetic calorimeter (ECAL) and the silicon tracker. A charged particle (muon) that is passing all sub-detectors is changing its trajectory as illustrated with a blue line. The presence of the magnetic field induced by the superconducting solenoid is noticed by the curvature of the different tracks. This figure is taken from Reference [15].

3.2.1 The coordinate system of CMS

In CMS, polar coordinates are used to describe the detector. The origin of the right-handed coordinate system is at the center of the detector. The z-axis points in the direction of the beam, the x-axis radially towards the center of the LHC, and the y-axis vertically upwards. The azimuth angle ϕ , the polar angle θ and the radius $r = \sqrt{x^2 + y^2}$ which represents the distance from the interaction point are those coordinates describing the CMS detector. The pseudorapidity

$$\eta = -\ln \tan\left(\frac{\theta}{2}\right) \quad (3.4)$$

corresponds to the longitudinal rapidity y for massless relativistic particles. Differences in rapidity are Lorentz-invariant under boosts along the z-direction, which is also the case for η for ultra-relativistic particles. The angular distance between particle i and j is given by

$$\Delta R(i, j) = \sqrt{(\Delta \eta(i, j))^2 + (\Delta \phi(i, j))^2}, \quad (3.5)$$

while one of the most important quantities in particle physics correspond to the transverse momentum

$$p_T = \sqrt{p_x^2 + p_y^2}, \quad (3.6)$$

respectively, of the detected particles. This quantity is also invariant under Lorentz boosts in the z-direction.

3.2.2 Magnet

When choosing the magnet, the CMS collaboration decided for a superconducting solenoid. It has to be cooled down to 4 K to ensure superconductivity. The solenoid is in total 12.5 m long, has a diameter of 6 m and a magnetic field strength of 3.8 T. The iron yokes surrounding the solenoid weigh over 10,000 tons altogether and are used to return the magnetic field and prevent its leakage outside the detector. The advantage of the compact construction of the CMS detector is that the calorimeters are within the magnetic field and only after them the iron yokes are distributed such that most of the particles (hadrons, electrons and photons) cannot interact with the iron yokes. The curvature of the magnetic field is used for the momentum measurement of charged particles.

3.2.3 Muon system

Since muons are minimum ionizing particles, they pass through ECAL and HCAL without being absorbed. Muons can be detected precisely, using the design of the CMS detector. The muon system is outside of the solenoid and is covering the pseudorapidity range of $|\eta| < 2.4$. It measures muon momenta using the Drift Tubes (DTs), Cathode Strip Chambers (CSCs) and Resistor Plate Chambers (RPCs) as shown in Fig. 3.3. The DTs are divided into four stations (Muon Barrel, MB1-4) and are used in the barrel range ($|\eta| < 1.2$) since the muon rate is lower there. Using these stations, both the position and the direction of muons are determined with a resolution of $100\mu\text{m}$ and 1mrad , respectively.

Compared to DTs, the response time of CSCs is faster which is why they are placed in the range of $1.2 < |\eta| < 2.4$ where the muon rate is higher (smaller angle to the beam, more particles). Like the DTs, the CSCs are divided into four stations (Muon endcap, ME1 - 4) which partially overlap in order to avoid gaps in acceptance. Muons can be reconstructed with an accuracy of $200\mu\text{m}$ in position and 10mrad in direction. The time response of RPCs is fast and with a resolution of 1cm they are used in the range of $|\eta| < 1.6$. More details can be found in Reference [14, 16].

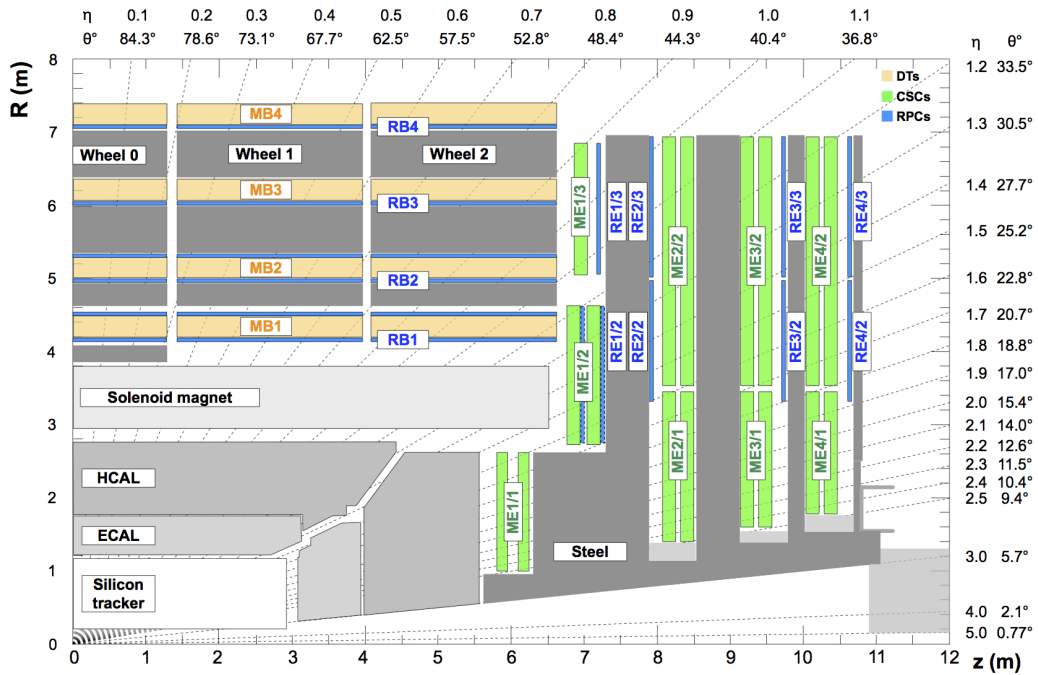


Fig. 3.3 A Schematic view of the muon system. The DTs (MB1 - 4) cover the pseudorapidity range $|\eta| < 1.2$, the CSCs (ME1 - 4) $1.2 < |\eta| < 2.4$ and the PRCs $|\eta| < 1.6$. This figure is taken from Reference [14].

3.2.4 Hadronic calorimeter

The HCAL consists of two different materials. For the absorbing material, brass and steel is used such that interactions lead to hadronic showers, which in turn interact with the next layer of towers, the plastic scintillators. The entire system is divided into four detector components. The barrel (HB), the endcap (HE), the outer (HO) and the forward (HF) region together cover a pseudorapidity range of $|\eta| < 5.2$. This sampling set of calorimeters is shown in Fig. 3.4. The barrel (HB) is constructed to stop most of the hadrons. It has a thickness which is around ten times higher than the interaction length of hadrons. It covers the pseudorapidity range of $|\eta| < 1.3$ and is arranged in so called towers (scintillator layers). The outer detector (HO) is covering the region $|\eta| < 1.26$ which is inside the MB discussed in Section 3.2.5 and also to see in Fig. 3.3. The endcap (HE) covers the pseudorapidity range of $1.3 < |\eta| < 3$ and the forward detector (HF) covers $|\eta| < 5.2$.

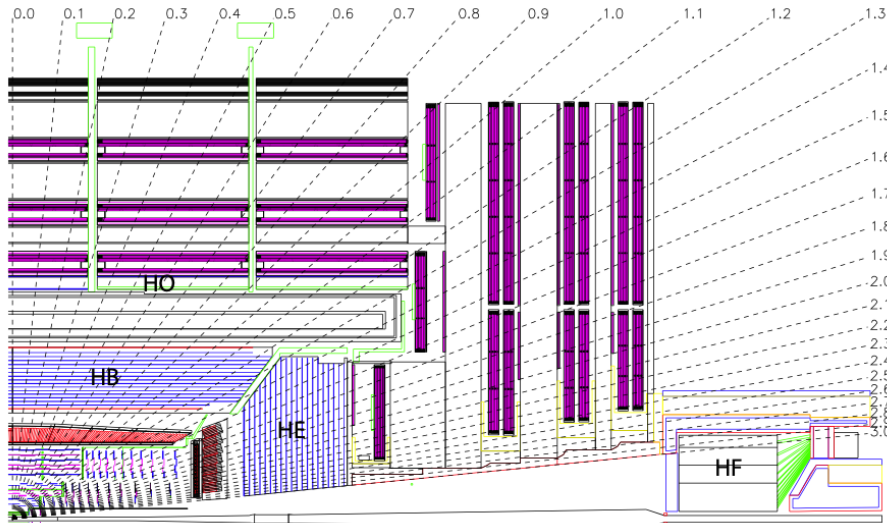


Fig. 3.4 A Schematic view of the HCAL. The entire system is divided into the barrel (HB, $|\eta| < 1.3$), the endcap (HE, $1.3 < |\eta| < 3$), the outer (HO, $|\eta| < 1.26$) and the forward (HF, $|\eta| < 5.2$) region. This figure is taken from Reference [12].

3.2.5 Electromagnetic calorimeter

The hermetic and homogeneous ECAL in CMS is made of lead tungstate (PbWO_4) crystals and is designed to detect electrons and photons due to energy deposits in the calorimeter. It is divided into two pseudorapidity regions. An illustration showing these different components of the ECAL is given in Fig. 3.5. The ECAL Barrel (EB) part covers the range $|\eta| < 1.479$ and the ECAL Endcap (EE) $1.479 < |\eta| < 3.0$. The latter has a preshower sensor, called

ECAL Sensor (ES), which is used to improve the position resolution of the detected photons. It is made of silicon and together the EE and EB absorb 98% of all electrons and photons.

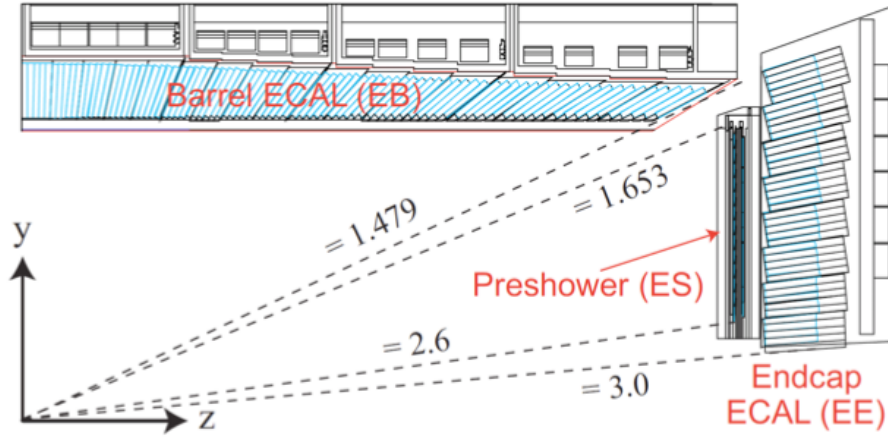


Fig. 3.5 A Schematic view of the ECAL system. The EB part covers the pseudorapidity range of $|\eta| < 1.479$ and the EE with the preshower sensor ES of $1.479 < |\eta| < 3.0$. This figure is taken from Reference [13].

3.3 The CMS Tracker and its Alignment

For this analysis, the most relevant part of the CMS detector is its heart, the world's largest silicon tracker. With the CMS tracker, trajectories of charged particles coming from the interaction point are reconstructed using the information of the position of hits (sensor signals) in the tracker modules. This section describes briefly the main aspects of it and its alignment while more details of it, its properties and alignment can be found in Reference [12].

3.3.1 The silicon tracker

The CMS tracker records hits³ resulting from the passage of charged particles. For a given bunch crossing, hits are used to reconstruct tracks and vertices. The vertex with tracks from particles with the highest energy in a collision is called the main primary vertex, explained in more detail in Section 4.1. In one bunch crossing typically not only two protons collide and the entire number of interactions is called pileup. Additionally, the tracker has to deal with displaced vertices, which are called secondary vertices and occur e.g. due to the production of hadrons containing a beauty valence quark. Due to their long lifetime ($10^{-12}s$) compared

³Hit is defined in Section 4.1.

to other hadrons, high- p_T c and b hadrons decay after some hundred μm . Secondary vertices are typically used for b-tagging as described in Section 4.4.2 in more detail. A schematic view of the entire tracker and its detector components is shown in Fig. 3.6.

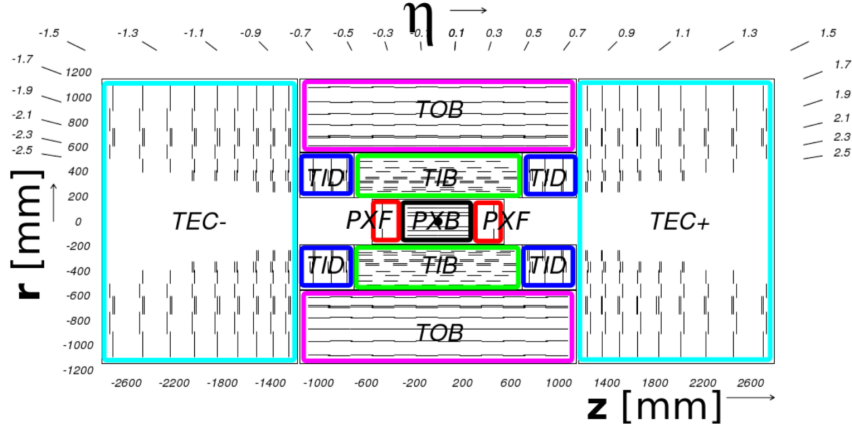


Fig. 3.6 A Schematic view of the tracker system and its detector components. The tracker is divided into four barrel pixel layers (BPix), three pixel endcap on disks each side (FPix), four tracker inner barrel layers (TIB), six outer tracker barrel layers (TOB), three tracker inner disk layers (TID) and nine tracker end caps layers (TEC). This figure is taken and modified from References [5, 17].

The tracker covers the region of $|\eta| < 2.5$, is 5.8m long, has a diameter of 2.5m and consists of two different types of detector modules, the *pixel* and *strip* detector. Compared to strip modules, pixel modules have a higher resolution. They are closest to the interaction point obtaining the reconstruction for primary and secondary vertices. The BPix layers are placed in the barrel region and the FPix endcap disks in the forward region. With the *Phase I* upgrade of the CMS detector, the pixel detector was replaced. The granularity of the pixel was increased, with four layers in BPIX and three discs in FPIX. Nowadays it consists of 1856 pixel modules. The *Phase I* upgrade was performed during the LHC shutdown in 2016/17 (EYETS⁴) to face higher luminosity during the years 2017 and 2018. A schematic view of the pixel detector is shown on the left for the *Phase 0* with two endcap disks and three barrel layers. On the right side in Fig. 3.7 is the comparison with the *Phase I* upgrade showing the barrel layers upgrade from three to four. The *strip* detector is divided into two regions, the tracker inner/outer barrel region (TIB and TOB) and the tracker inner/endcap region (TID and TEC). Since the particle density decreases with an increase in distance from the interaction point, the density of placed modules decreases in the same way. The tracker strip detector consists of 15148 strip modules while the entire silicon tracker includes 66 million pixels and 10 million strip sensors.

⁴Extended Year End Technical Stop

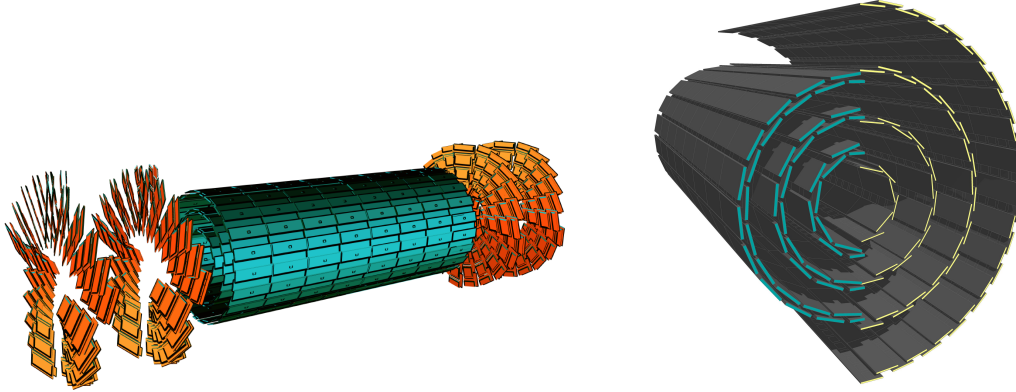


Fig. 3.7 Left, *Phase-0* with three barrel layers. Right, *Phase I* upgrade showing the barrel layers upgrade from three to four. These figures are taken from References [18, 19].

The silicon tracker has to be aligned since the precision of the installation of the modules does not correspond to their resolution. Additionally, systematic misalignments can occur leading to a systematic bias in the reconstruction. The alignment procedure has to take this into account, since a systematic bias in the reconstruction can lead to a systematic bias in the measurement.

3.3.2 Tracker alignment

This section gives a brief introduction to track-based alignment⁵ [20, 21] and shows the study of so-called weak modes, which was part of the technical task for this thesis. The challenge of the alignment procedure is illustrated in Fig. 3.8. The alignment procedure has to deal with $\mathcal{O}(100k)$ shifts, rotation and curvatures to evaluate such that residuals are minimized.

In total, the picture looks even more complicated. Many tracks can be distorted and even missing hits shown here as a missing blue dot in Fig. 3.9 are the task for the alignment algorithm. The track based alignment algorithm, which is known as MillePede, has the challenge to minimize the residuals between the measured and expected hit positions in the detector. It uses a least square approach to minimize the sum of the squares of normalized residuals which is given by

$$\chi^2(p, q) = \sum_j^{\text{tracks}} \sum_i^{\text{hits}} \left(\overbrace{\frac{m_{ij} - f_{ij}(p, q_j)}{\sigma_{ij}}}^{\text{residual}} \right)^2. \quad (3.7)$$

⁵The tracker can be aligned with laser too.

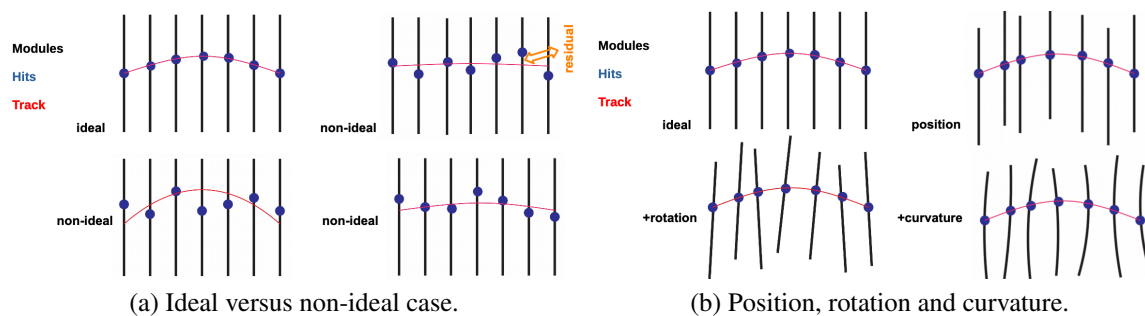


Fig. 3.8 An illustration of the challenge of the alignment. In a) the different scenarios which show an ideal and non-ideal case. In b) it can be seen with what the alignment procedure had to deal with it per scenario. Modules are shown as black strokes, hits as blue dots and the lines in red correspond to the tracks.

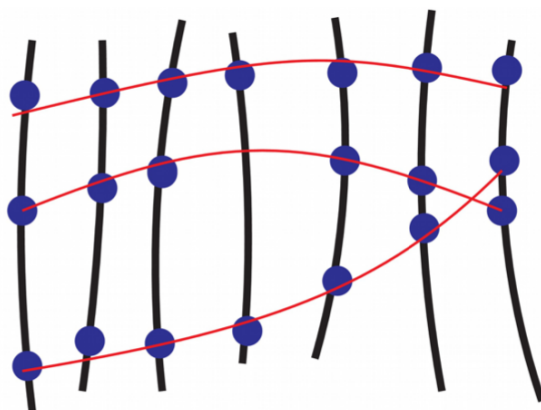


Fig. 3.9 Modules are black, hits blue and tracks red. A missing hit is represented as a missing blue dot here. In total, MillePede has to calculate over 200 thousand parameters that occur due to the number of different tracks, the curvatures, rotations and shifts in the position of the modules.

MillePede solves then the resulting large equation system by using a global-fit approach with the linearisation of Eq. 3.7. The output is a number of geometry parameters corresponding to the positions of the modules. In this way, modules are aligned using a track-based alignment procedure by the minimization of the residuals between the measured and expected hit positions in the detector.

A comparison of an ideal with a realistic set of geometry parameters is shown in Fig. 3.10 on the left (right) before and (after) the track based alignment procedure. By the comparison of right and left it can be seen that the residuals are minimized after this procedure.

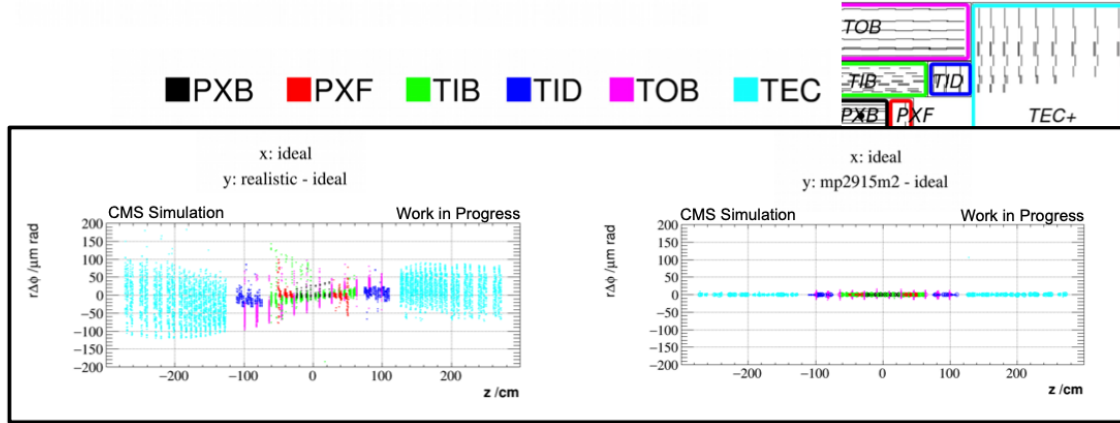


Fig. 3.10 An illustration of the movement of the modules in $r\Delta\Phi$ is shown as a function of z . Left, the difference between a realistic and an ideal geometry, on the right after the alignment procedure.

Weak modes

Global distortions to which Eq. 3.7 is not sensitive are called weak modes. The track-hit residuals shown in Fig. 3.8 are not affected by these. They can lead to a systematic bias in the measurement. Weak modes can only be avoided by additional information, such as including data and physics knowledge in Eq. 3.7. To investigate those potential weak modes a systematic study is needed. A full list of weak modes is shown in Appendix A, while in the following one of them is described briefly.

The weak mode *Twist* occurs as the name already hints due to a tiny twist of the entire detector. The impact on the tracks of this weak mode and how it can be treated is illustrated in Fig. 3.11. On the one hand, for the weak mode study $Z^0 \rightarrow \mu^+ \mu^-$ decays are used, since the mass of that boson is well known. On the other hand, cosmic rays should not be bent in the detector without the presence of a magnetic field. Both approaches need to be and were studied within this thesis to ensure the precision of the silicon tracker in CMS. One of the results is shown in Fig. 3.12. It shows the reconstruction of data including a misalignment induced by the weak mode twist in green and the result after the alignment procedure by using the mass constraint of $Z^0 \rightarrow \mu^+ \mu^-$ events in red. A comparison with an ideal alignment case in blue is also shown.

The impact on physics of different alignments

Two different alignment eras for 2018 are studied to test whether there is an impact on reconstructed secondary vertices due to the improvement of the alignment. Two approaches were made, one where the selected datasets have a small difference in the alignment and

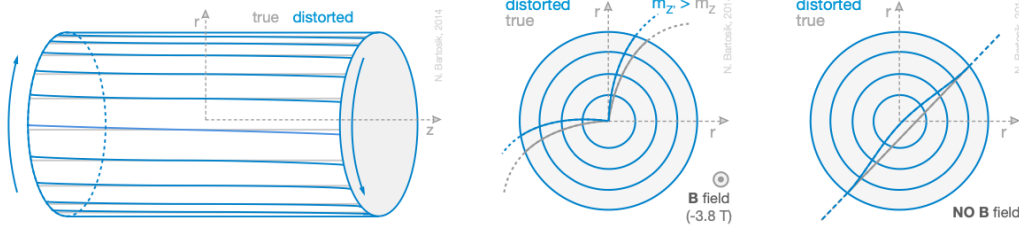


Fig. 3.11 An illustration of the weak mode twist at two different perspectives. This figure shows the distorted and true positions of the tracks from $Z^0 \rightarrow \mu^+ \mu^-$ under the impact of a magnet field strength of zero and 3.8 T. This figure is taken from Reference [22].

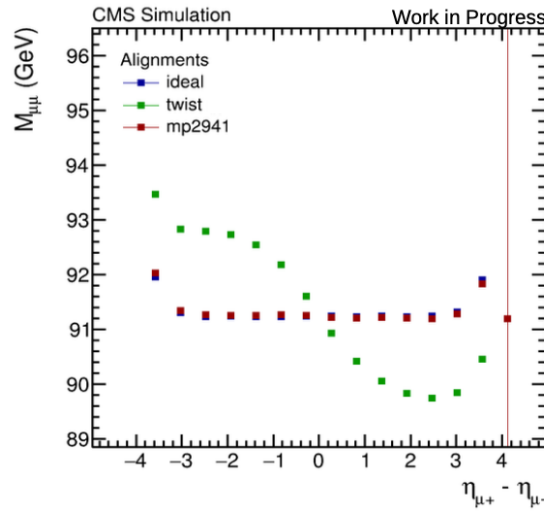


Fig. 3.12 Three different alignment scenarios are shown here. In blue is the ideal case, in green an alignment including the weak mode *Twist* and in red the result after the alignment procedure.

the other where the difference is large. One data set with a prompt reconstruction⁶ and for the same runs with the same events one data set with the latest so called UL (ultra legacy⁷) reconstruction were used. For this purpose, a Charmonium data set is selected and two analyses are performed. One is a D^* meson analysis, the topic of this study which is explained in detail in Chapter 6, and the other is a J/Ψ analysis. The latter was selected since the statistics is higher for the process $J/\Psi \rightarrow \mu^- \mu^+$ and the dedicated data set, respectively. The entire study can be found in Reference [23]. The main results are shown in the following only for the subset where the differences are large between the different alignment areas.

⁶Prompt reconstruction is performed during data taking.

⁷Ultra legacy (UL) is used in CMS to name the latest reconstruction for a dedicated data set.

The method for the signal extraction for D^* mesons is shown in Fig. 3.13 for a data set with a prompt (UL) reconstruction on the left (right). The obtained signal number of $N_{sub}(D^{*\pm})$ is here reflecting that the impact on physics due to improved alignment is small.

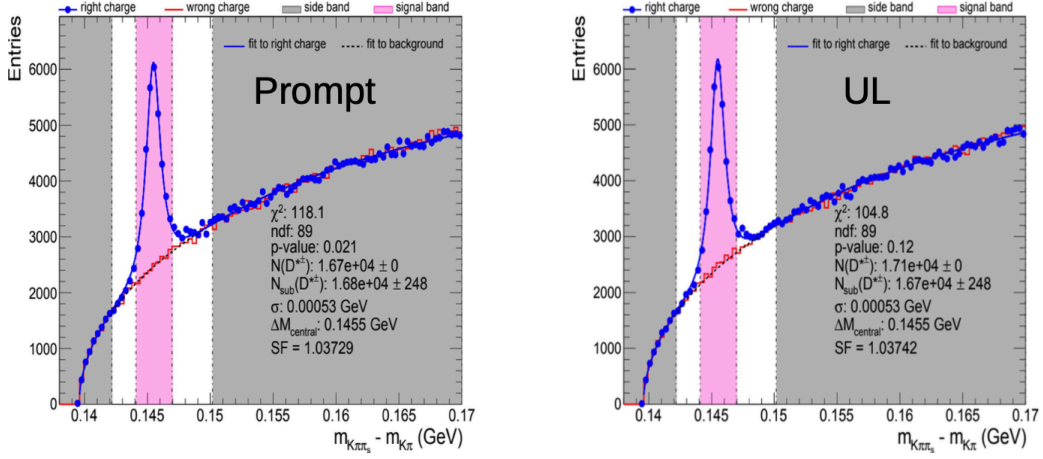


Fig. 3.13 Comparison of the impact of two different alignments (prompt reco versus UL) to a D meson analysis. On the left is the signal extraction for D meson for prompt reconstruction and on the left is for the ultra legacy processing.

The reconstructed J/Ψ mass distributions in the signal region for an invariant mass of $\in [2.95, 3.25]$ GeV are shown in Fig. 3.14. With the improvement of the alignment a better mass resolution is obtained and also the distribution for the decay length is narrower. In the case of the prompt reconstruction alignment the distribution is shifted to the right. The entire study showed that for the kinematic range of the reconstructed particles and the studied distributions the impact on the physics due to the improved alignment is small for D^* mesons, and in the case of the J/Ψ an impact of the improved alignment is visible. More comparisons can be found in the Appendix A.

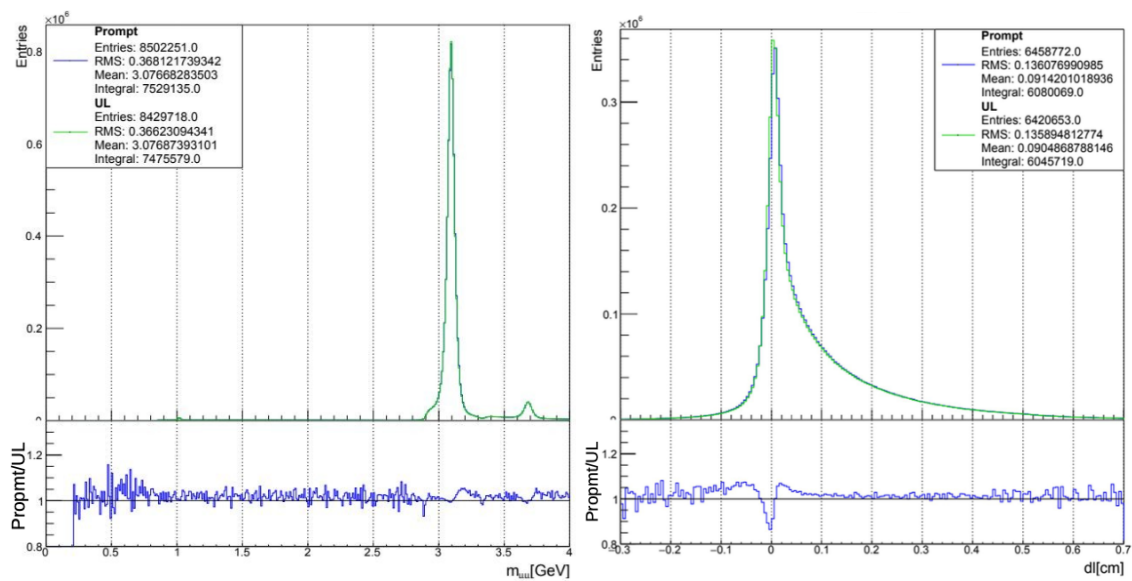


Fig. 3.14 Comparison of the impact of two different alignments (prompt reco vs UL) on a J/Ψ analysis where the difference between the two alignment areas is large. The distribution in green belongs to the prompt reco alignment and in blue to the UL alignment, On the left, is the invariant mass and on the right the decay length for J/Ψ 's.

References

- [1] P. M. Watkins, “DISCOVERY OF THE W AND Z BOSONS”, *Contemp. Phys.* **27** (1986) 291–324, 10.1080/00107518608211015.
- [2] “Tim Berners-Lee, World Wide Web inventor”, <https://cds.cern.ch/record/39437>.
- [3] ATLAS Collaboration, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”, *Phys. Lett.* **B716** (2012) 1–29, 10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
- [4] CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, *Phys. Lett.* **B716** (2012) 30–61, 10.1016/j.physletb.2012.08.021, arXiv:1207.7235.
- [5] CMS Collaboration, “The CMS experiment at the CERN LHC”, *Journal of Instrumentation* **3** (aug, 2008) S08004–S08004, 10.1088/1748-0221/3/08/s08004.
- [6] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider”, *Journal of Instrumentation* **3** (aug, 2008) S08003–S08003, 10.1088/1748-0221/3/08/s08003.
- [7] ALICE Collaboration, “The ALICE experiment at the CERN LHC”, *Journal of Instrumentation* **3** (aug, 2008) S08002–S08002, 10.1088/1748-0221/3/08/s08002.
- [8] LHCb Collaboration, “The LHCb Detector at the LHC”, *Journal of Instrumentation* **3** (aug, 2008) S08005–S08005, 10.1088/1748-0221/3/08/s08005.
- [9] CMS Collaboration, “CMS Luminosity Measurements for the 2016 Data Taking Period”, technical report, CERN, Geneva, 2017. <https://cds.cern.ch/record/2257069>.
- [10] CMS Collaboration, “CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV”, technical report, CERN, Geneva, 2019. <https://cds.cern.ch/record/2676164>.
- [11] E. Mobs, “The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019”, <https://cds.cern.ch/record/2684277>. General Photo.
- [12] CMS Collaboration, “The CMS Experiment at the CERN LHC”, *JINST* **3** (2008) S08004, 10.1088/1748-0221/3/08/S08004.
- [13] CMS Collaboration, “The CMS electromagnetic calorimeter project”. CMS Technical Design Report. CERN, Geneva, 1997. [Addendum: CERN-LHCC-2002-027].
- [14] CMS Collaboration, “Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s}=13$ TeV”, *Journal of Instrumentation* **13** (jun, 2018) P06015–P06015, 10.1088/1748-0221/13/06/p06015.
- [15] D. Barney, “CMS Detector Slice”, technical report, Jan, 2016. <https://cds.cern.ch/record/2120661>, CMS Collection.

- [16] CMS Collaboration, “Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV”, *JINST* **7** (2012) P10002, 10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.
- [17] “TWiki page: CMS Tracker Alignment Performance Results 2018”, https://twiki.cern.ch/twiki/pub/CMSPublic/TkAlignmentPerformanceMid18/Figures_Experimental_Apparatus_Tracker_colours.png.
- [18] CMS Collaboration, “Commissioning and performance of the CMS pixel tracker with cosmic ray muons. Commissioning and Performance of the CMS Pixel Tracker with Cosmic Ray Muons”, *JINST* **5** (Nov, 2009) T03007. 37 p, 10.1088/1748-0221/5/03/T03007, arXiv:0911.5434.
- [19] A. Kumar, “Near future upgrades of the CMS pixel detector”, *Journal of Instrumentation* **10** (2015).
- [20] C. Kleinwort, “General broken lines as advanced track fitting method”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **673** (2012) 107–110, <https://doi.org/10.1016/j.nima.2012.01.024>.
- [21] CMS Collaboration, “Alignment of the CMS tracker with LHC and cosmic ray data”, *Journal of Instrumentation* **9** (jun, 2014) P06009–P06009, 10.1088/1748-0221/9/06/p06009.
- [22] N. Bartosik, “Associated Top-Quark-Pair and b-Jet Production in the Dilepton Channel at $\sqrt{s} = 8$ TeV as Test of QCD and Background to tt +Higgs Production”, Aug, 2015. Presented 03 Jul 2015. <https://cds.cern.ch/record/2047049>.
- [23] “Testing the Alignment on benefit for Physics - J.M.”, https://indico.cern.ch/event/992964/contributions/4340548/attachments/2234857/3788158/AlignmentBenefitForPhysics_V4.pdf.

Chapter 4

Event Reconstruction and Data Acquisition

Contents

4.1	Tracks and Vertices	50
4.2	Muons and Electrons	51
4.3	Particle Flow	52
4.4	Jets	53
4.4.1	Clustering	53
4.4.2	b-jet identification	54
4.5	Missing Transverse Energy	54
4.6	Acquisition and Analysis of Data	55
4.6.1	Trigger system	56
4.6.2	Data formats and analysis	57
4.7	MC and Detector Simulations	58
	References	61

Events recorded by CMS consist of several objects which need to be reconstructed for an analysis. This chapter provides an overview of the reconstruction of particles and, if possible, their tracks. The purpose is to summarize briefly the reconstruction of the main physical objects entering most of the analyses in CMS. The study in this thesis relies especially on the reconstruction of tracks, primary and secondary vertices.

4.1 Tracks and Vertices

In CMS, tracks of all charged particles passing through the tracker in an event are reconstructed via the identification of hits in the tracker [1]. Hits are charge deposits and are combined to tracks under the presence of a strong magnetic field. The momentum of particles is measured from the reconstruction of the trajectory, usually a helix. Due to a typically large number of tracks in an event, a technique is needed to assign the multiple hits correctly to tracks.

For the track reconstruction, an iterative Combinatorial Track Finder (CTF) algorithm is used. The CTF algorithm starts to identify hits in the tracker as tracks and to analyse these iteratively, beginning with high momentum tracks which are next to the collision point. It forms tracks beginning from the inside of the tracker where a higher granularity exists due to the pixel detectors in contrast to the outer part of the tracker where the strip detector is placed. It continues by removing these hits, which were used to form high momentum tracks, and reconstructing the more difficult tracks with a lower momentum. Its procedure is based on a Kalman filter [2, 3] and starts iteratively with the track reconstruction with a small collection of hits called seeds. An extrapolation of these seed trajectories is performed to the outside of the detector, and additional hits are analysed under the assumption of helical paths. The reconstruction is performed taking into account the energy loss of the particles due to scattering and radiation while traversing the detector. If the extra hits are compatible with the seed trajectories, the tracking candidate is saved in a collection for further checks. The Kalman filter performs a χ^2 fit of all track candidates and keeps only "good" ones, based on a minimum number of hits in the tracker and the goodness of fit. In this way, all tracks are reconstructed, and the best track candidates are determined.

The interaction point of incoming particles or decay point of outgoing particles is called a vertex. These vertices are represented with orange dots in the event display shown in Fig. 4.1. Primary vertices (PV) correspond to the interaction points of the collision of one of the large number (10^{11}) of colliding protons inside a bunch from each side. Typically there is more than one PV in an event at the crossing of the proton bunches at the LHC. To find these vertices, all reconstructed tracks are analysed, and the vertex reconstruction [1] starts to reconstruct all vertices. The reconstruction of primary vertices focuses on the tracks with a low χ^2 and small impact parameter. The procedure continues by clustering these tracks according to the adaptive vertex fitter [4], and deterministic annealing [5]. The reconstruction of secondary vertices is performed in the same way with the additional condition that the particles impact parameter coming from secondary vertices can be larger compared to the one from particles coming from primary vertices.

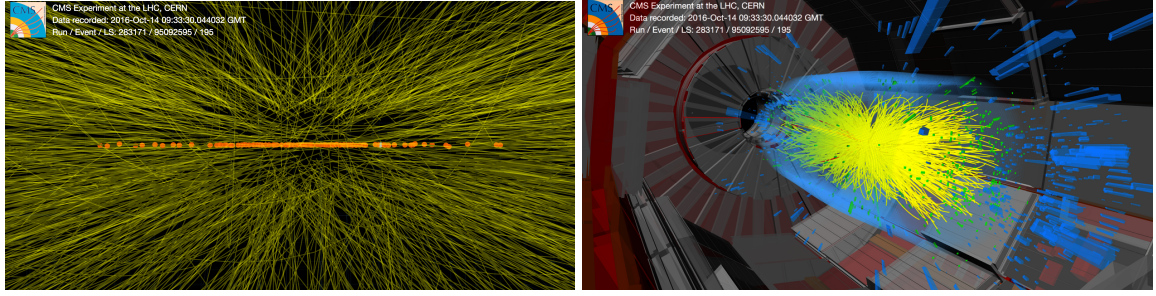


Fig. 4.1 On the left, a zoomed view of an event of proton-proton collisions in the center of the CMS detector (tracker) can be seen. The reconstructed vertices are presented with orange dots on the horizontal axis while the lines correspond to the reconstructed tracks. Here, the scale is equal to a few centimeters. Each track corresponds to a charged particle, and every one of these tracks must be associated with only one vertex. On the right, the same event of the proton-proton collisions shows signals in regions of the CMS detector beyond the tracker. These figures are taken from Reference [6].

4.2 Muons and Electrons

Muons are minimum ionizing particles, and their reconstruction is performed in CMS by an initially independent reconstruction of tracks in the tracker and muon chambers [7, 8]. It can be distinguished between global and tracker muons.

The algorithm for the reconstruction of global muons starts to reconstruct tracks in the muon chambers and extrapolates these to the inner tracks of the detector, the tracker. This approach uses the Kalman filter [3]. It performs a global fit with the information about the hits in the muon chamber and the tracker. The reconstruction is performed by starting with hits in the muon chambers and combining the information of the sub-detectors. Thus, the momentum resolution for global muons for high p_T muons ($p_T > 200$ GeV) is improved compared to tracker muons.

In contrast to global muons, the reconstruction of tracker muons starts from the tracker. Tracks are extrapolated from the tracker to the muon chambers and are matched with hits in the muon chambers without fitting them. Multiple scattering from muons traversing the whole detector is taken into account. In this way even, by using single hits in the muon chambers, the reconstruction for tracker muons for lower p_T muons is more efficient compared to global muons.

For **electrons** the reconstruction procedure is different. By traversing the detector material under the presence of the strong magnetic field they lose energy due to bremsstrahlung. The resulting change in the curvature and their energy loss have to be taken into account during the reconstruction of the electron track candidates. Their energy deposit is measured in the

ECAL, and the reconstruction of electrons starts from there. Energy deposits from energy clusters, and those which are next to each other are clustered into a super cluster (SC). The energy of the SC corresponds to the sum of the energy of all neighboring clusters, while the weighted mean of them determines the electrons position in the ECAL. Either electron seeds are identified by starting extrapolating electron track candidates from the ECAL clusters or by extrapolating them from the tracker to the ECAL. Since the emission of photons can create electron-positron pairs, the algorithm for the electron reconstruction is more complex compared to the one for muons. While it fits track from the beam spot to the surface of the ECAL, it has to take the curvature and bremsstrahlung ambiguities from electrons into account. A more detailed discussion of the reconstruction of electrons can be found in Reference [9].

4.3 Particle Flow

The idea of particle flow (PF) reconstruction is to make use of the combination of the information of all sub-detectors which in turn can lead to a better momentum resolution of the reconstructed objects. Such a reconstruction is called PF event reconstruction as illustrated in Fig. 4.2. In the following, a short explanation of it is given while more details can be found in Reference [10]. Particle flow candidates are photons, muons, electrons, charged, and neutral hadrons.

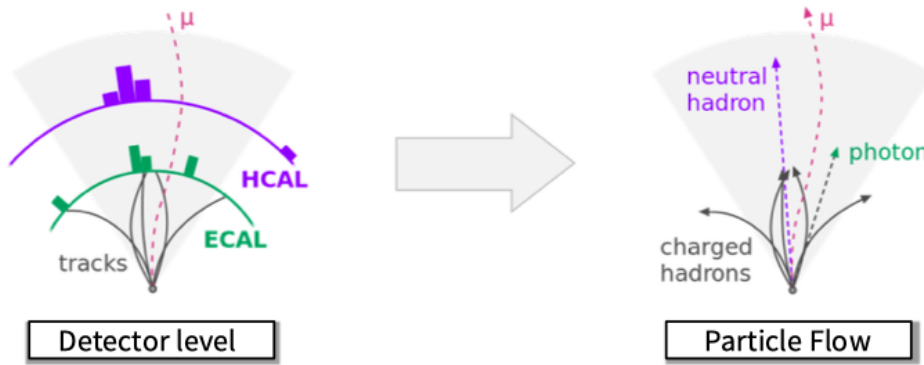


Fig. 4.2 This figure illustrates the particle flow reconstruction. On the left (right), it can be seen how an event looks without (with) PF reconstruction. This figure is taken from Reference [11] and modified.

Photons can not be tracked with the tracker except photon conversions. Their reconstruction is performed via the measurement in the ECAL and relies always on the combination of the information of all sub-detectors.

Muons are reconstructed by combining the energy deposits in the muon system with hits from global muons. Their energy is determined via the measurement of their curvature through the entire detector.

Particle flow **electrons** are reconstructed by matching their track to the energy deposit in the ECAL as discussed in Section 4.2. The energy loss due to bremsstrahlung is taken into account in this procedure. The energy of PF electrons is determined with a combinatorial measurement of the bremsstrahlung, energy deposits in the ECAL, and the momentum of the electron.

Charged hadrons can leave signals in all sub-detectors except in the muon chambers. The determination of their energy is a combination of the measurement of the sub-detectors (tracker, ECAL, HCAL). **Neutral hadrons** leave no signals in the tracker; their energy is determined only by the ECAL and HCAL. In both cases, the measured energies for hadrons have to be corrected since detector-specific effects like the response function of the calorimeters need to be taken into account for the measurement.

With particle flow event reconstruction, jets can be clustered, and other event properties like the missing transverse energy can be determined.

4.4 Jets

Due to the QCD confinement, colored particles (partons) cannot be detected directly but only as bound states (hadrons). The measurement of their kinematic properties, e.g., the p_T of partons, can only be achieved via the reconstruction of the produced bunch of hadrons resulting from them, which forms a jet. A brief discussion about jet reconstruction is given in the following, while a more detailed description of the reconstruction of jets, the jet energy correction, and the b-jet identification can be found in References [12–14].

4.4.1 Clustering

In CMS, a lot of analyses use jets clustered from PF candidates. By default, the reconstruction of clustered jets is performed via the anti- k_t algorithm [15]. The measured jet energy has to be corrected. This correction aims to determine the momentum vector of a jet such that it corresponds almost to the energy of the parton from which the jet originates. Since partons are not free particles, their transverse momentum measurement is not directly measurable. To obtain reasonable predictions via simulated events, the entire jet momentum is therefore compared with jets on the particle level. Jets are clustered at particle level from all (with the

exception of neutrinos) stable particles ($c\tau > 1$ cm) and matched with simulated reconstructed PF jets.

4.4.2 b-jet identification

Processes like the associated Higgs production can be filtered from other processes based on the physical properties of b quarks and b-flavored hadrons, respectively. B-hadrons contain a b-quark, are the heaviest hadrons with a rest mass of more than 5 GeV and decay suppressed via the weak interaction into charm hadrons, one of which contains a valence c-quark. Due to these properties, b hadrons have a long lifetime ($\tau \approx 0.5$ mm/c), and beauty decays produce an electron or muon in about 20% of the decays. B-hadrons carry a large part of the entire jet momentum since, within the fragmentation of b-quarks to hadrons, most of the energy is given to the b-hadron. Therefore, they can have a high boost and an even longer effective lifetime compared to other hadrons. This can shift the vertex of the beauty decay by several mm from the primary vertex. Leptons (hadrons too) produced via beauty decays leave a displaced track in the tracker. This signature is used for the identification of b-jets. An illustration of b-jet identification is shown in Fig. 4.3.

The algorithms of the b-jet identification are implemented in b-taggers. These days, there are several b-taggers, for instance, the DeepJet jet-flavor identification algorithm [16]. In CMS, all jet algorithms have in common that the reconstruction of jets works most efficiently above a certain threshold¹. Jet algorithms can cluster jets less well near the production threshold of the underlying beauty production process compared to high momentum jets. Within this thesis, this fundamental physical problem is avoided, and beauty events are tagged even at the production threshold via the identification of a D^* coming from a b-hadron as explained in detail in Chapter 6.

4.5 Missing Transverse Energy

The presence of neutrinos or neutral BSM² particles that travel through the detector without leaving any signal like tracks in the tracker or energy deposits in the calorimeters can only be identified via the missing transverse momentum. This is typically called missing transverse energy $\vec{E}_T$ (MET) and can be determined by the use of momentum conservation. The sum of the transverse momenta of all reconstructed particles has to be zero after the interaction of the colliding particles since they carried no transverse momentum before the interaction. A

¹In CMS the working point is at ≥ 20 GeV

²Beyond Standard Model

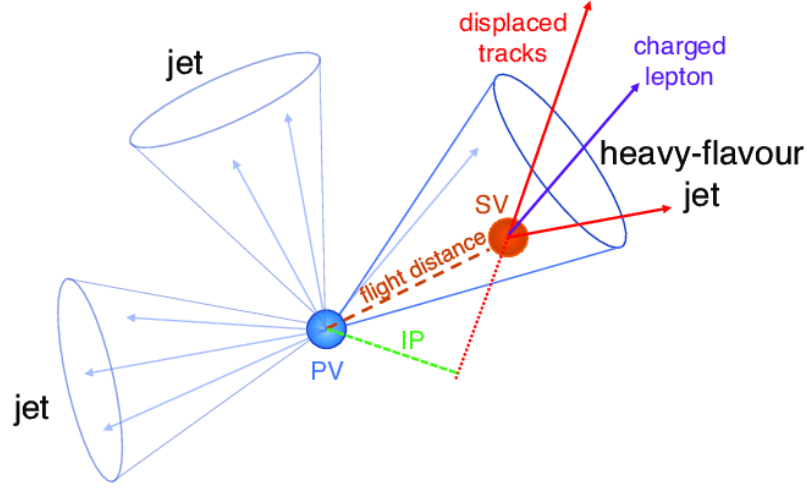


Fig. 4.3 Illustration of a heavy-flavor jet reconstruction with the use of the SV. At the SV displaced tracks and a jet results from the charmed or beauty hadron decay. Due to the long lifetime of b-hadrons, the produced particles appear displaced from the PV with a large impact parameter (IP). This figure is taken from Reference [14].

global description of the missing transverse energy is achieved by the use of the combined information of all sub-detectors, including all PF objects. It is given by

$$\vec{E}_T = -\sum_i \vec{p}_{Ti} \quad (4.1)$$

where the sum of momenta in the transverse plane of all PF objects corresponds to the negative sum of the missing transverse energy. A more precise discussion can be found in Reference [17].

4.6 Acquisition and Analysis of Data

A collision rate of 40 MHz is achieved at the LHC, whereby it is impossible to record every event since an average event size corresponds to 1 MB. To not process every event, particular criteria are used to decide whether the event is recorded. These criteria are manifested in a system called trigger system [18]. Only a small fraction of all collisions are then processed and recorded for offline analysis. The online selection of data in CMS is performed by a two-stage trigger system, the *Level 1* (L1) and *High Level Trigger* (HLT).

4.6.1 Trigger system

The collision rate to be processed at the LHC is reduced from 40 MHz to 100 kHz by the L1 trigger. The L1 trigger decides within $3.2 \mu s$ whether the event is passed to the HLT or not [19]. Since such a decision has to be made quickly, only simple checks can be made at this level. For example, it is checked whether there is a narrow shower in the ECAL, which indicates photons or electrons. Since Run 2 it is also possible to check whether the photon or electron is isolated. In addition, global information can be used to search for high-energy candidates in the event [20]. A sketch of the architecture of the L1 trigger with all its sub-components is shown in Fig. 4.4.

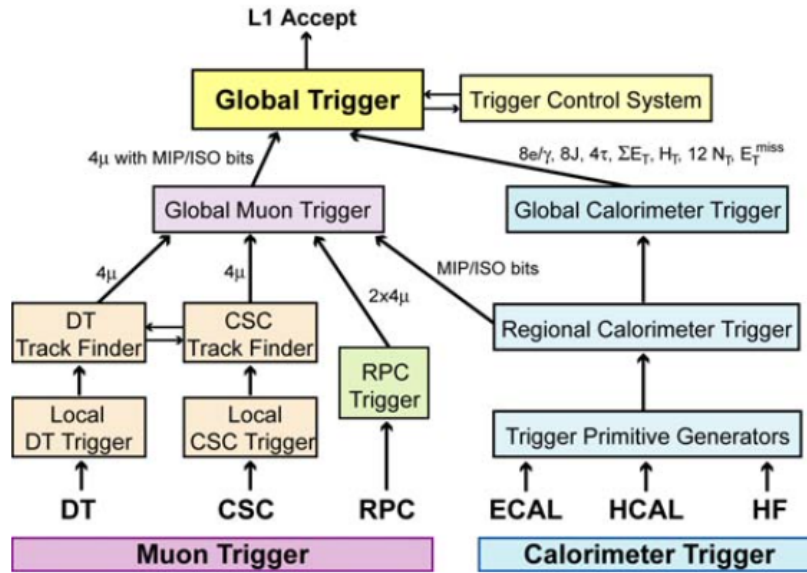


Fig. 4.4 Illustration of the architecture of the L1 trigger system. This figure is taken from Reference [21].

After running through the L1 trigger, the trigger system reduces the collision rate to be processed to 1 kHz such that dedicated events can be recorded and their data stored. This is done with the help of the HLT, which consists of several trigger paths [22]. Algorithms are implemented in each of these trigger paths, which reconstruct the objects and are based on criteria according to which the event will be recorded or not. First, it is checked whether simple criteria are fulfilled, such as checking if an activity is present in the calorimeter. Then more complicated criteria are checked, which then take place at the level of the track reconstruction. There are plenty of trigger paths which in general, are optimized by each physics analysis group (PAG) for dedicated analyses, such as, for instance, a Higgs boson study. In this thesis, charm and beauty production is studied, which corresponds approximately to 10% and 1% of the total pp cross section, respectively. In many cases, there

are no specific triggers for these kinds of low p_T studies, which is why in this study, the core total-charm and total-beauty cross-section analysis is based on the Minimum Bias (MB) path responsible for the recording of MB-like datasets. MB triggers react on generic detector signals, which are mostly in the forward region, triggered by the remnant of inelastic pp collisions. These events are then collected on MB datasets, but, worth mentioning, despite their name, these datasets often also contain other triggers, which are biasing the events, e.g., the vertex multiplicity or track multiplicity. The data sets used in this work were only recorded with true MB triggers. These are listed in Section 6.1.1. The used minimum bias collisions for this analysis are discussed in Section 6.1 and listed in Tab. 6.1.

4.6.2 Data formats and analysis

Events accepted by the trigger system contain the full recorded information from the detector. This format is called RAW data, and it is classified into several distinct primary datasets, which are stored in different Tier centers. The reconstruction process is with the detector simulation in CMS the most CPU-intensive activity. Reconstructed (RECO) data is produced at the Tier 0/1 center. This already contains most of the high-level objects used in analyses together with low-level information. However, this format is too large to be used efficiently for analyses and thus only stored at the Tier-0 center. The tiering system starts at 0 and ends at 3, where centers with a lower tier value provide a large number of computing resources and the associated necessary storage space [23]. There is only one Tier 0 center located at CERN, several Tier 1 centers, and hundreds of Tier 2 and Tier 3 centers located at institutes worldwide, such as the Tier 2 center at DESY. The data flow between the different Tier centers is shown in Fig. 4.5.

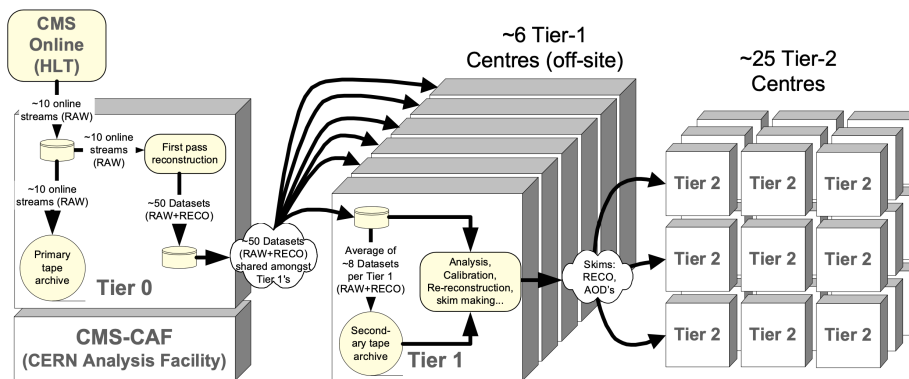


Fig. 4.5 Data flow between the different Tier centers. This figure is taken from Reference [21].

The RECO format is then reduced to the Analysis Object Data (AOD) format, which is stored at centers with a Tier value ≥ 1 . The AOD size is about 1 MB per event and contains all relevant information such as reconstructed high-level objects (tracks, electrons, muons, and jets) with which almost all physical analyses can be performed. A new format called miniAOD [24], whose size is only 50 kB per event, was introduced with Run 2 at the CMS. Due to a preselection, this format contains a reduced amount of information about the reconstructed high-level physics objects. The format NANOAO [25] requires only ≈ 1 KB per event. Like the miniAOD, it does not contain as much low-level information as the AOD and has even less information than the miniAOD. However, with NANOAO, analyses can be performed which only require high-level information. All data formats are reconstructed and analyzed with software based on C++, called CMSSW (CMS SoftWare). The NANOAO format can also be analyzed with plain ROOT [26, 27] and C++, respectively.

4.7 MC and Detector Simulations

Partonic final states of hard scattering processes for fixed-order calculations in perturbative QCD are not directly experimentally observable. A connection from the partonic final state to the observed hadronic particles must be established. As mentioned in Section 2.3.3, to simulate QCD processes, Monte Carlo (MC) event generators were designed, such that they can be used for analyses at different energy scales of physical processes [28, 29]. Those simulated events are then passed through a detector simulation to be comparable to data, which is necessary for experimental analyses. For instance, via generator events, an optimization of the event selection without experimental bias or a description of the detector responses can be achieved while the detector response can be simulated via passing the MC simulation through a virtual detector simulation followed by a processing through the data acquisition system for the reconstruction of data.

Interactions of the final states with the detector material and magnetic field, leading to the production of secondary particles, particle decay, Bremsstrahlung, energy loss, or scattering, need to be taken into account within this detector simulation. For detector simulation the toolkit GEANT4 [30] is typically used in CMS. The idea is to perform the same analysis on the Monte Carlo data like on real experimental data. In this way, MC events are fully simulated, passed through the reconstruction chain, and can then be used to determine the detector efficiency for a certain process, for instance, $pp \rightarrow BX \rightarrow D^*X$. The detector efficiency is a necessary ingredient for the cross-section measurements. It is determined in this thesis for the studied processes and enters the measurement described in Chapter 6.

Scattering events simulated by MC event generators make use of parton shower (PS) models, which obey the same QCD factorization theorem used for the PDF evolution described in Section 2.3. First, the matrix element of the hard scattering process is obtained from fixed-order calculations, and then the PS takes place. Via a procedure, which takes non-perturbative approaches into account [31], partonic final states are then evolved to hadrons. The emissions of partons, both in the initial and final state, are described via the DGLAP equations till a certain energy scale is reached. An illustration of an event for a typical hadron-hadron collision simulated with the MC technique is shown in Fig. 4.6. For the hadronization, hadron decays or soft processes, like multiple parton interactions, phenomenological models are necessary to describe the observed phenomena. One of the hadronization models is for instance the Lund string model used by PYTHIA6 [32, 33] which is explained in more detail in Chapter 2.3.

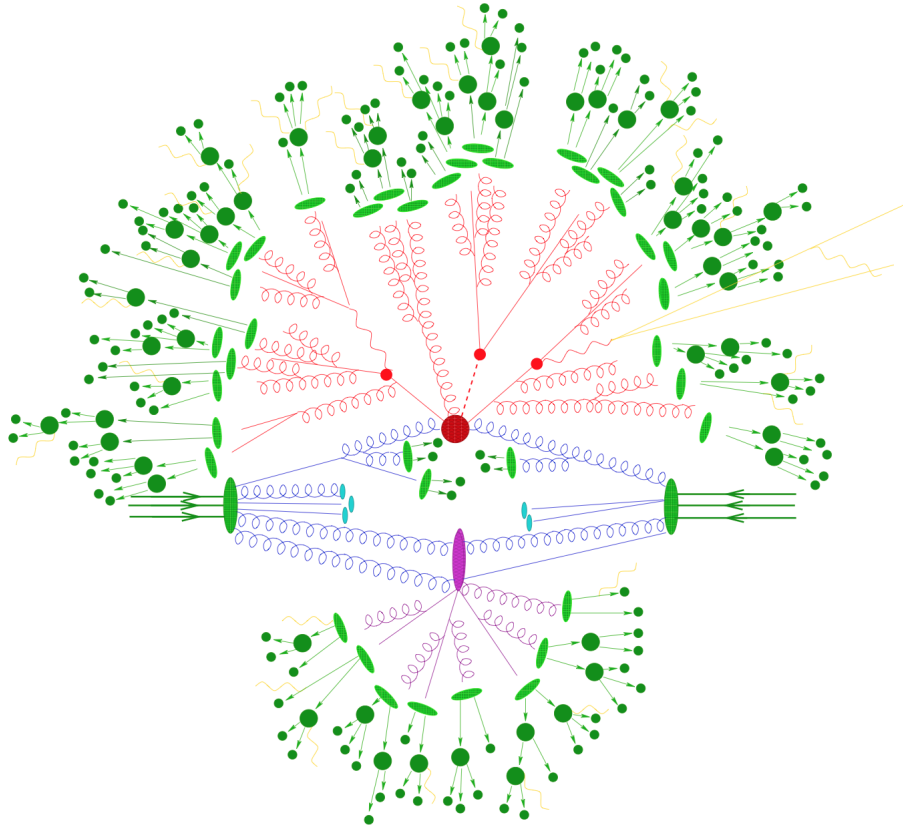


Fig. 4.6 A typical event of a hadron-hadron collision simulated by a Monte-Carlo event generator. The hard collision in the center (red blob) is surrounded by the underlying event. Bremsstrahlung is simulated by the PS. The sketch shows a secondary hard scattering event (purple blob), hadronization (green light blobs), hadron decays (dark green blobs) and soft photon radiation (yellow lines). This figure is taken from Reference [34].

Within this procedure, a matching³ between the calculation of the matrix element and the added PS is needed, which prevents double counting of real emissions. PS simulate not only soft collinear emissions but semi-hard emissions too. The calculation of the matrix element describes hard emissions (large angle) more accurately than the PS, and thus, this matching tackles the correction of the hard emissions from the PS. In this way, both the fixed-order calculation and the feature of the PS (leading-logarithmic resummation) are ensured. The MC generated with PYTHIA6 is at the level of LO+PS and is used in Chapter 6 for shape comparison and the determination of the detector efficiency. While PYTHIA6 is at the level of LO+PS, newer particle physics event generators like POWHEG [35] are at the level of NLO+PS, but they are not designed for heavy quark processes at low energy scales. Attempts were made within this study, but several parameters need to be tuned for this purpose, which is a difficult application. Worth mentioning, for instance, the scale parameter h_{damp} ⁴, which declares at which energy scale the PS takes place, is by default far away from the kinematic range of the low- p_T heavy quark processes studied in this thesis, since these MC generators are designed for high- p_T processes like Higgs production. Thus, a full NLO+NLL (next-to-leading-logarithmic) calculation is used for the theoretical QCD predictions of the cross sections of heavy quark flavor processes at lower energy scales, while PYTHIA6 is used to obtain efficiency corrections. Details of the Fixed-Order-next-to-leading-logarithmic (FONLL) theory predictions for heavy quark flavor processes [36, 37] used in this work are described in Chapter 5.

³A matching is not performed for a LO calculation.

⁴The damping parameter h_{damp} is also known as resummation scale

References

- [1] CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker”, *Journal of Instrumentation* **9** (oct, 2014) P10009–P10009, 10.1088/1748-0221/9/10/p10009.
- [2] W. Adam, B. Mangano, T. Speer, and T. Todorov, “Track Reconstruction in the CMS tracker”, technical report, CERN, Geneva, Dec, 2006.
<https://cds.cern.ch/record/934067>.
- [3] R. Frühwirth, “Application of Kalman filtering to track and vertex fitting”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **262** (1987), no. 2, 444–450,
[https://doi.org/10.1016/0168-9002\(87\)90887-4](https://doi.org/10.1016/0168-9002(87)90887-4).
- [4] R. Frühwirth, W. Waltenberger, and P. Vanlaer, “Adaptive vertex fitting”, *J. Phys. G* **34** (2007) N343, 10.1088/0954-3899/34/12/N01.
- [5] K. Rose, “Deterministic annealing for clustering, compression, classification, regression, and related optimization problems”, *IEEE Proc.* **86** (1998), no. 11, 2210–2239, 10.1109/5.726788.
- [6] CMS Collaboration, “Collisions recorded by the CMS detector on 14 Oct 2016 during the high pile-up fill”, (Nov, 2016). CMS Collection.
- [7] CMS Collaboration, “Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s}=13$ TeV”, *Journal of Instrumentation* **13** (jun, 2018) P06015–P06015, 10.1088/1748-0221/13/06/p06015.
- [8] CMS Collaboration, “Performance of CMS Muon Reconstruction in pp Collision Events at $\sqrt{s} = 7$ TeV”, *JINST* **7** (2012) P10002, 10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.
- [9] CMS Collaboration, “Performance of Electron Reconstruction and Selection with the CMS detector in proton-proton collisions at $\sqrt{s}=8$ TeV”, *Journal of Instrumentation* **10** (jun, 2015) P06005–P06005, 10.1088/1748-0221/10/06/p06005.
- [10] CMS Collaboration, “Particle-flow reconstruction and global event description with the CMS detector”, *JINST* **12** (2017), no. 10, P10003, 10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.
- [11] CMS Collaboration, “PF concept in CMS.”,
http://bartosik.pp.ua/hepsketches/cms_particle_flow.
- [12] G. P. Salam, “Towards jetography”, *The European Physical Journal C* **67** (2010), no. 3, 637–686, 10.1140/epjc/s10052-010-1314-6.
- [13] CMS Collaboration, “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV”, *JINST* **12** (2017), no. 02, P02014, 10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.

- [14] CMS Collaboration, “Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV”, *JINST* **13** (2018), no. 05, P05011, 10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.
- [15] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_t jet clustering algorithm”, *JHEP* **04** (2008) 063, 10.1088/1126-6708/2008/04/063, arXiv:0802.1189.
- [16] E. Bols et al., “Jet flavour classification using DeepJet”, *Journal of Instrumentation* **15** (dec, 2020) P12012–P12012, 10.1088/1748-0221/15/12/p12012.
- [17] CMS Collaboration, “Performance of missing energy reconstruction in 13 TeV pp collision data using the CMS detector”, technical report, CERN, Geneva, 2016. <https://cds.cern.ch/record/2205284>.
- [18] CMS Collaboration, “The CMS trigger system”, *JINST* **12** (2017), no. 01, P01020, 10.1088/1748-0221/12/01/P01020, arXiv:1609.02366.
- [19] CMS Collaboration, S. Dasu et al., “CMS. The TriDAS project. Technical design report, vol. 1: The trigger systems”, 12, 2000. CERN-LHCC-2000-038.
- [20] CMS Collaboration, “CMS Technical Design Report for the Level-1 Trigger Upgrade”, technical report, Jun, 2013. <https://cds.cern.ch/record/1556311>.
- [21] CMS Collaboration, “The CMS experiment at the CERN LHC”, 10.1088/1748-0221/3/08/s08004.
- [22] CMS Collaboration, “CMS: The TriDAS project. Technical design report, Vol. 2: Data acquisition and high-level trigger”,.
- [23] C. Grandiet al., “CMS Computing Model: The "CMS Computing Model RTAG"", technical report, CERN, Geneva, Dec, 2004. <https://cds.cern.ch/record/814248>.
- [24] G. Petrucciani, A. Rizzi, and C. Vuosalo, “Mini-AOD: A New Analysis Data Format for CMS”, 10.1088/1742-6596/664/7/072052.
- [25] CMS Collaboration, “A further reduction in CMS event data for analysis: the NANO AOD format”, *EPJ Web Conf.* **214** (2019) 06021. 6 p, 10.1051/epjconf/201921406021.
- [26] R. Brun and F. Rademakers, “ROOT - An Object Oriented Data Analysis Framework”, *Nucl. Instrum. Meth. A* **389** (1997) 81–86, 10.1016/S0168-9002(97)00048-X.
- [27] R. Brun et al., “root-project/root: v6.18/02”, August, 2019. 10.5281/zenodo.3895860.
- [28] S. Weinzierl, “Introduction to Monte Carlo methods”, arXiv:hep-ph/0006269.
- [29] A. Buckley et al., “General-purpose event generators for LHC physics”, *Phys. Rept.* **504** (2011) 145–233, 10.1016/j.physrep.2011.03.005, arXiv:1101.2599.
- [30] S. Agostinelli et al., “Geant4—a simulation toolkit”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003), no. 3, 250 – 303, [https://doi.org/10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8).

- [31] P. Nason, “A New method for combining NLO QCD with shower Monte Carlo algorithms”, *JHEP* **11** (2004) 040, 10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.
- [32] T. Sjostrand, L. Lonnblad, and S. Mrenna, “PYTHIA 6.2: Physics and manual”, arXiv:hep-ph/0108264.
- [33] T. Sjostrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual”, *JHEP* **05** (2006) 026, 10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.
- [34] S. Höche, “Introduction to parton-shower event generators”, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders*, pp. 235–295. 2015. arXiv:1411.4085. 10.1142/9789814678766_0005.
- [35] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”, *JHEP* **11** (2007) 070, 10.1088/1126-6708/2007/11/070, arXiv:0709.2092.
- [36] M. Cacciari, M. Greco, and P. Nason, “The P(T) spectrum in heavy flavor hadroproduction”, *JHEP* **05** (1998) 007, 10.1088/1126-6708/1998/05/007, arXiv:hep-ph/9803400.
- [37] M. Cacciari, S. Frixione, and P. Nason, “The p(T) spectrum in heavy flavor photoproduction”, *JHEP* **03** (2001) 006, 10.1088/1126-6708/2001/03/006, arXiv:hep-ph/0102134.

Chapter 5

Physics of Heavy Quark Flavor Production

Contents

5.1 Heavy Flavor Production and Decays in pp Collisions	65
5.2 Theory Predictions for Heavy Flavor Production	68
5.3 Existing Measurements of Charm and Beauty Production	71
References	81

This chapter provides a brief description of the physics of heavy quark (without the top quark) production in pp collisions. It shows a selection of relevant existing results of measurements of D meson production in different detectors at different center of mass energies. The theoretical aspects and the use of theory predictions for heavy quark flavor production are covered in this chapter. The aim of this chapter is to provide all necessary terminology and context for the results of this analysis presented in Chapter 6.

5.1 Heavy Flavor Production and Decays in pp Collisions

In pp collisions at the LHC, the dominating process for heavy flavor production is the gluon fusion process. As an example, Feynman diagrams for beauty production at LO are shown in Fig. 5.1. For charm production, they look the same (replacing beauty with charm).

As explained in Chapter 2, the quarks are not free particles but do form hadrons. Charm quarks hadronize into charmed hadrons like D mesons, while beauty quarks hadronize into b hadrons. The probability that a quark hadronizes into a certain hadron $\mathcal{H}$ carrying a momentum fraction z is described by the fragmentation function, which enters the QCD

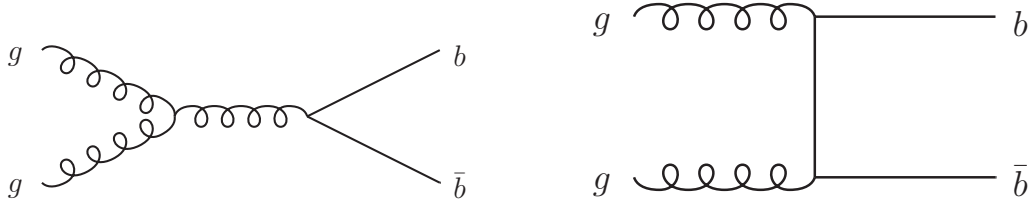


Fig. 5.1 Feynman diagrams for beauty production at LO. On the left the s channel and on the right the t channel.

cross section calculation c.f. Eq. 2.39. This function is similar to the PDF which describes the beginning of QCD process, while the fragmentation function describes the end of a QCD process. Different parameterizations exist for fragmentation functions explained in more detail in Section 5.2. The normalization factor of the fragmentation function is called fragmentation fraction. It corresponds to the integral of the fragmentation function, i.e. the probability that a certain quark hadronizes into a certain hadron. The average values of the measured fragmentation fractions based on LEP¹ [1] measurements are listed in Fig. 5.1 on the left for charm. On the right, these values correspond not to a fragmentation fraction but to a fraction describing the hadronization of a beauty quark into B hadron which decays into a charmed hadron $\mathcal{H}_c$. This measurement [1] used the known branching fraction measurement from BELLE² [3] and CLEO³ [4]. In the case of beauty decay, an anti-charm and strange

Table 5.1 Values for the fragmentation fractions $f(c \rightarrow \mathcal{H}_c)$ and the fraction and decay value $f(b \rightarrow \mathcal{H}_c)$ are derived from the LEP measurements [1]. The index c or b corresponds to charm or beauty, respectively.

$\mathcal{H}_c$	$f(c \rightarrow \mathcal{H}_c) [\%]$	$f(b \rightarrow \mathcal{H}_c) [\%]$
D^0	$54.2 \pm 2.4 \pm 0.7$	$58.7 \pm 2.1 \pm 0.8$
D^+	$22.5 \pm 1.0 \pm 0.5$	$22.3 \pm 1.1 \pm 0.5$
D_s^+	$9.2 \pm 0.8 \pm 0.5$	$13.8 \pm 0.9 \pm 0.6$
D^{*+}	$23.6 \pm 0.6 \pm 0.3$	$22.1 \pm 0.9 \pm 0.3$
Λ_c^+	$5.7 \pm 0.6 \pm 0.3$	$7.3 \pm 0.8 \pm 0.4$

quark can be produced in addition by a W^- as shown in Fig. 5.2.

When charm or beauty hadrons are produced in pp collisions, they can fly several hundred μm or more depending on their lifetime and Lorentz boost. The decay lengths for D mesons

¹The Large Electron Positron Collider at CERN.

²B detector at Koo Energy Ken (KEK) at the High Energy Research Accelerator Organization in Tsukuba, Japan. Acronym taken from Fermilab news [2].

³CLEO is a detector at the Cornell Electron Storage Ring which is a high-luminosity electron-positron collider at the Wilson Synchrotron Laboratory, Cornell University

and some b hadrons, their different masses and their different quark contents are listed in Tab. 5.2.

Table 5.2 List of properties of charm and beauty hadrons from PDG 2020 [5]. The mass, the quark content and the decay length of the individual hadron is shown.

Hadron	Quark content	Mass [MeV]	Decay length $c\tau[\mu\text{m}]$
B^+	$u\bar{b}$	5279.34 ± 0.12	491.1
B^0	$d\bar{b}$	5279.65 ± 0.12	455.4
B_s^0	$s\bar{b}$	5366.88 ± 0.14	454.2
Λ_b^0	udb	5619.60 ± 0.17	441.0
D^+	$c\bar{d}$	1869.60 ± 0.16	311.8
D^0	$c\bar{u}$	1864.83 ± 0.14	122.9
D_s^+	$c\bar{s}$	1968.47 ± 0.33	149.9
Λ_c^+	udc	2286.46 ± 0.14	60.7
D^{*+}	$c\bar{d}$	2010.25 ± 0.14	

The measurement of this work deals with D^* meson decays. In this thesis, D^* mesons coming from a b hadron decay are non-prompt and those who do not are prompt D^* mesons. At hadron (meson) level the production of a non-prompt D^* is shown in Fig. 5.2. Here, the

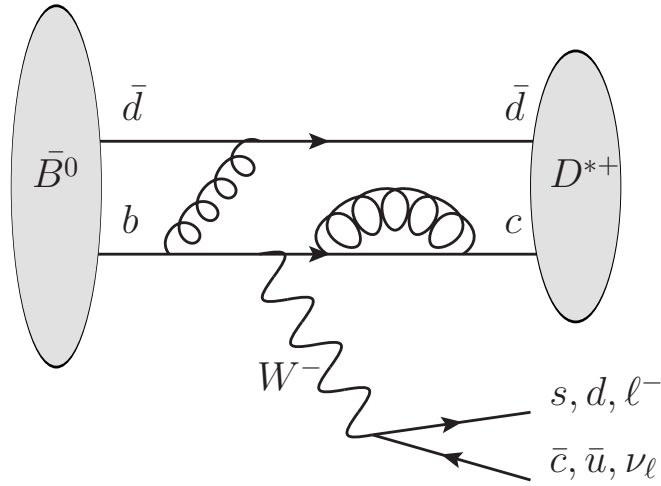


Fig. 5.2 Decay of $B \rightarrow D^*X$ ($X = s\bar{c}, d\bar{u}, \ell\nu_\ell$) at meson level.

$\bar{B}^0$ is decaying into a D^{*+} and some remnant X corresponding here to a quark pair ($X = s\bar{c}$). This non-prompt D^* can then decay into different other mesons. A decay of a prompt D^* into a D^0 and π_s (s=slow) is shown at meson level in Fig. 5.3a.

Here the index s corresponds to slow i.e. the π_s carries very little momentum, since in this process the remaining phase space (≈ 145 MeV, c.f. Tab. 5.2) for it corresponds almost

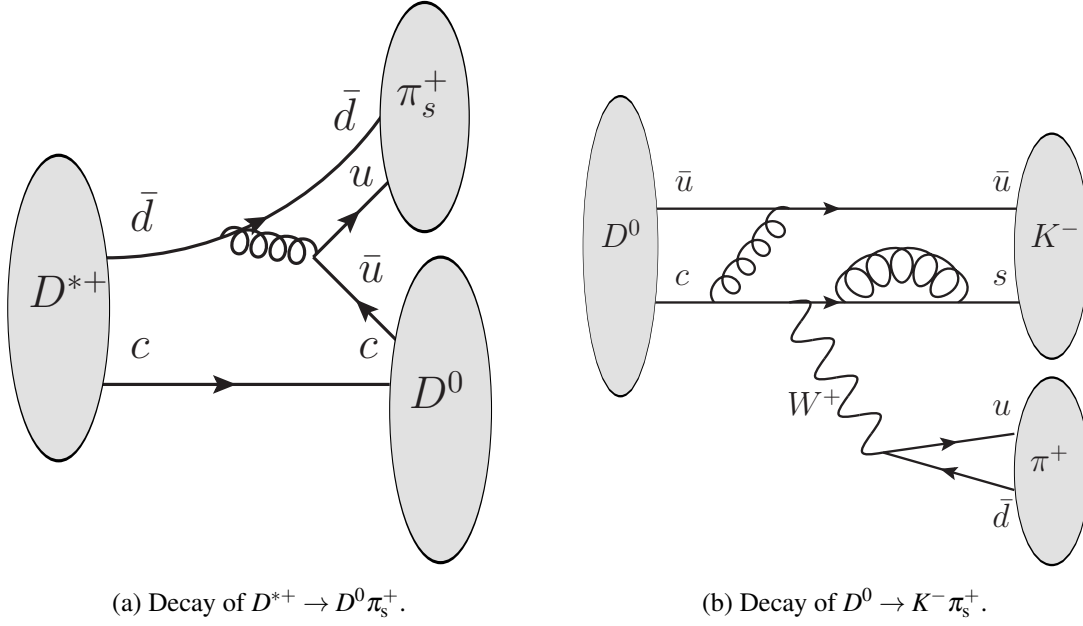


Fig. 5.3 Decay of $D^{*+} \rightarrow D^0 \pi_s^+$ and of $D^0 \rightarrow K^- \pi_s^+$. These mesons are prompt D mesons.

to the mass of the π . The decay branching fractions $\mathcal{B}$ for this specific D^* decay and for the D^0 decay used are shown in Tab. 5.3.

Table 5.3 List of branching fractions taken from PDG 2020 [5].

$\mathcal{B}(D^0 \rightarrow K^- \pi^+)$	$(3.95 \pm 0.03)\%$
$\mathcal{B}(D^{*\pm} \rightarrow D^0 \pi^\pm)$	$(64.7 \pm 0.9)\%$
$\mathcal{B}(D^{*\pm} \rightarrow D^0 \pi^\pm \rightarrow (K^- \pi^+) \pi^\pm)$	$(2.67 \pm 0.03)\%$

A decay of a prompt D^0 into a kaon (K) and a π is presented in Fig. 5.3b. Here the π can carry higher momentum compared to a π_s in Fig. 5.3a. This analysis relies on the reconstruction of D^* mesons decaying into a D^0 and slow pion π_s , while the D^0 is decaying into a K and π . This entire process is represented in Fig. 5.4.

5.2 Theory Predictions for Heavy Flavor Production

This section is used to describe the theory predictions for heavy quark flavor production in pp collisions used in this work. It is divided into first, a brief description of the MC generator PYTHIA6 [6, 7]. This generator is used for the shape comparison with the measurements of inclusive, prompt and non-prompt D^* production and the determination of the corresponding

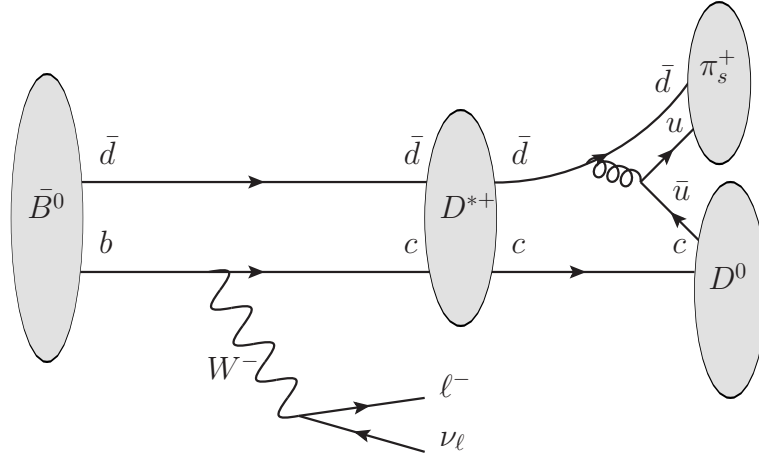


Fig. 5.4 Decay chain of $B \rightarrow D^* X \rightarrow D^0 \pi_s X \rightarrow K \pi \pi_s X (X = \ell \nu_\ell)$ at meson level.

efficiencies and acceptances. Secondly, the theory calculation FONLL which is used for the theory predictions for prompt and non-prompt D^* production in this work is described.

MC simulations are discussed in Section 4.7 in general, while in the following the description is reduced to the context of the analysis of this thesis. Since in this work, charm and beauty processes are measured close to their production threshold region, Pythia6 (version 424) is used to generate events with soft QCD processes with no requirement on the minimum or maximum on p_T . For the purpose of the charm beauty separation explained in Chapter 6, a $D^0 \rightarrow K\pi$ MC 5 TeV sample with a TuneZ2star tune is used. The tune uses the Bowler [8] fragmentation function for heavy flavor within the hadronization procedure. Details about the MC sample can be found in Section 6.1.2.

For heavy quark flavor production, theory predictions in this thesis correspond to the FONLL (Fixed Order Next-to-Leading Log) [9–12] calculation implemented via the publicly available FONLL web interface [13]. Via this interface, the predictions of heavy flavor production at FONLL approximation are obtained in this work. The main parameters used in the FONLL web interface for the theory prediction of charm and beauty production in pp collisions are discussed in the following.

Predictions for single inclusive distributions for a particle P are obtained as explained in [13] as a numerical convolution. The latter consists of a perturbative cross section $d\sigma_Q^{FONLL}$, a non-perturbative fragmentation function $D_{Q \rightarrow H_Q}^{NP}$ and a decay function $g_{H_Q \rightarrow P}^{\text{decay}}$, which describes that a hadron is decaying into a particle P . The equation is given by [9]

$$d\sigma_P^{FONLL} = d\sigma_Q^{FONLL} \otimes D_{Q \rightarrow H_Q}^{NP} \otimes g_{H_Q \rightarrow P}^{\text{decay}}. \quad (5.1)$$

Here, for a heavy flavor hadron, for instance the charmed hadron, the integral of $D_{Q \rightarrow H_Q}^{NP}$ corresponds to the fragmentation fraction for charm $f(c \rightarrow \mathcal{H}_c)$ shown in Tab. 5.1. The decay spectra $g_{H_Q \rightarrow P}^{\text{decay}}$ and the parameters of the non-perturbative (NP) fragmentation function $D_{Q \rightarrow H_Q}^{NP}$ are best determined from e^+e^- data [9].

The non perturbative fragmentation function $D_{Q \rightarrow H_Q}^{NP}$ for predictions for prompt D^* production is BCFY⁴ [14] parameterized and the parameters are obtained by a fit to BELLE and CLEO D^* data [14]. More information about fragmentation functions can be found for instance in References [15, 16]. The PDF CTEQ6.6 [17] is used here, which is the default PDF at 5 TeV on the FONLL web interface. The uncertainty of the PDF is summed in quadrature to scale and mass uncertainties. In the case of charm production, the fragmentation factor for the hadronization of a charm quark into a D^* of $f_c = 0.236$ c.f. Tab. 5.1 is used. For the mass, a central value of $m_c = 1.5$ GeV with the mass uncertainty of $m_c = 1.3, 1.7$ GeV is used. This uncertainty is summed in quadrature to the scale uncertainties [13]. The scale uncertainties μ_F and μ_R are discussed in Section 2.3.2 and how an uncertainty band in this work is derived is illustrated in Fig. 2.8. The central values for these theory predictions are obtained by $\mu_R = \mu_F = \mu_0 = \sqrt{m^2 + p_T^2}$, where in this context m stands for the quark mass. The lower and upper values of the uncertainty band are obtained by [9]

$$\Delta_{\pm} = \sqrt{\Delta_{\pm, \text{scales}}^2 + \Delta_{\pm, \text{mass}}^2 + \Delta_{\pm, \text{PDF}}^2}. \quad (5.2)$$

In case for beauty, the $b \rightarrow B$ prediction is derived with FONLL by parameterizing the non-perturbative fragmentation function $D_{Q \rightarrow H_Q}^{NP}$ in Eq. 5.1 with the Kartvelishvili function [18] with the parameter $\alpha = 24.2$. Here no fragmentation fraction value is applied since the fraction $b \rightarrow B$ (all B hadrons) corresponds to 1. To obtain non-prompt D^* production predictions, the FONLL prediction is convoluted with a $B \rightarrow D^*$ decay table extracted from Pythia6. This conversion table is shown in Fig. 5.5. Finally the fraction value of $f_b = 0.218$ is applied. As mentioned in Section 5.1, this corresponds not to the fragmentation value for beauty quarks but rather to the fraction describing a beauty quark hadronization in a b hadron convoluted with the decay fraction for the decay $B \rightarrow D^*$. For the PDF setup, CTEQ6.6 [17] is used and for the beauty mass a central value of $m_b = 4.75$ GeV with the mass uncertainty $m_b = 4.5, 5.0$ GeV is used. The mass and PDF uncertainties are summed in quadrature to scale uncertainties.

⁴Braaten, Cheung, Fleming & Yuan

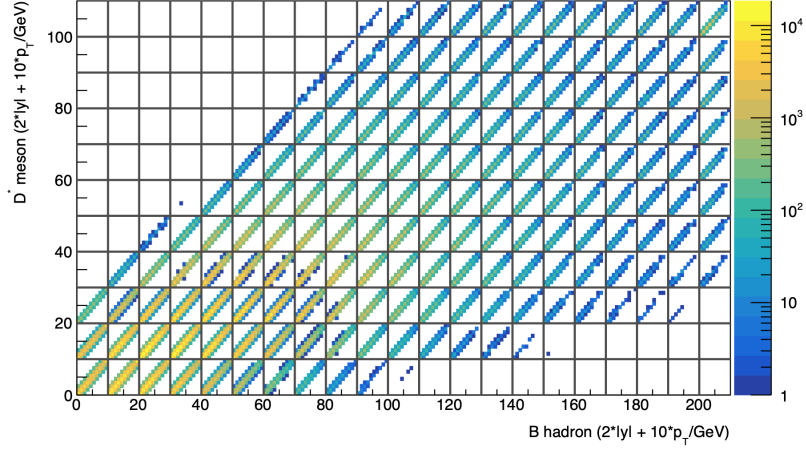


Fig. 5.5 $B \rightarrow D^*$ conversion table for FONLL extracted by Pythia6. This figure is taken from Reference [19].

5.3 Existing Measurements of Charm and Beauty Production

To connect the results of this thesis to the context of the existing measurements of charm and beauty production in pp collisions, a selection of relevant existing measurements is shown and discussed briefly in this section. A selection of measurements of prompt D mesons at different center of mass energies is given briefly, while more can be found in Reference [20]. The focus is on the description of those results of the measurements of prompt, and where possible, non-prompt D meson production at 5 TeV covering the phase space presented in Fig. 5.6. Before, the discussion starts with a measurement of inclusive D^* production at 7 TeV by ATLAS [25], which is their only measurement of D meson production. In this work a measurement of inclusive D^* production is performed 5 TeV.

ATLAS

The measurement of inclusive D^* production at 7 TeV by ATLAS [25] is performed differentially in p_T . The single differential measurement is performed in pseudorapidity up to $|\eta|(D^{*\pm}) < 2.1$ and presented in Fig. 5.7. Here, the p_T range starts at 3.5 GeV. The results are compared with predictions of different event generators some of which are described in Section 4.7. Additionally, a comparison with a FONLL prediction is shown, which is described in Section 5.2 in more detail. The ratio in Fig. 5.7 shows differences between the predictions and data. Here no separation of prompt from non-prompt D^* is performed.

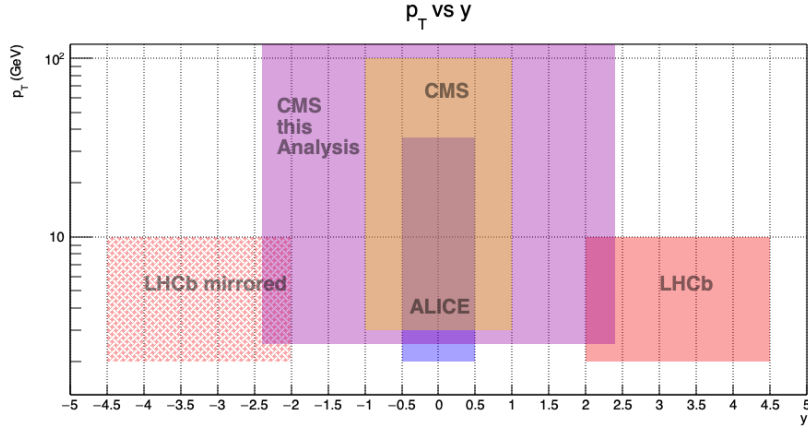


Fig. 5.6 Example for the covered phase space of the measurement of prompt D meson production at 5 TeV. In orange the measurement for prompt D^* production by ALICE [21] and in red the measurement by LHCb [22] for D^0 and D^* production. The measurement from CMS prompt [23] D^0 and non-prompt [24] D^0 production is shown in dark purple and the D^* production measurement of this work in light purple.

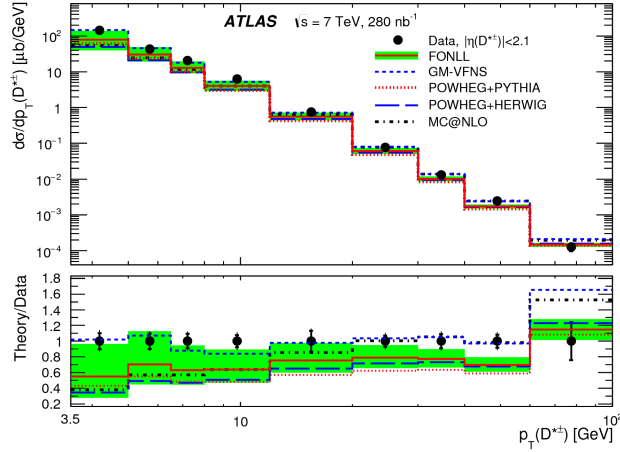


Fig. 5.7 Differential cross sections for $D^{*\pm}$ mesons as a function of p_T for data (points) compared to the NLO QCD calculations of FONLL, POWHEG+PYTHIA, POWHEG+HERWIG, MC@NLO and GM-VFNS (histograms). The data points are drawn in the bin centres. The inner error bars show the statistical uncertainties and the outer error bars show the statistical and systematic uncertainties added in quadrature. Uncertainties linked with the luminosity measurement are not included in the shown systematic uncertainties. The bands show the estimated theoretical uncertainty of the FONLL calculation. This measurement is performed at 7 TeV and the figure including its description is taken from ATLAS [25].

CMS

- Within a PhD thesis [26] cross sections for prompt D^* production are measured within the CMS experiment at 7 TeV using a very similar approach as the one in this thesis apart from the charm beauty separation. The double differential measurement is performed in an equidistant binning in p_T with a bin width of 1 GeV and in rapidity $|y|$ with a bin width of 0.5. The last bin in p_T represents the overflow bin. The results are presented in Fig. 5.8.

The measurements starts at 1 GeV in p_T such that almost the entire threshold region of the measured production process is covered. In this analysis, the fraction of the charm contribution in D^* final states is assumed to be 0.90 ± 0.05 [26]. A comparison with FONLL predictions for prompt D^* is given in each bin in the covered phase space. Additionally, a comparison with the results of ALICE [27] in the central rapidity range ($|y| < 0.5$) is shown and with LHCb in $2 < |y| < 2.5$.

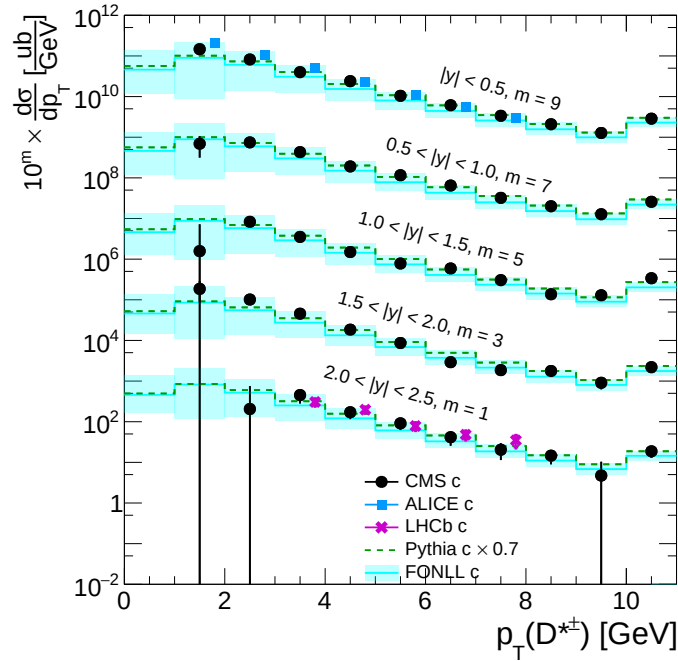


Fig. 5.8 Measurement of double differential cross section as a function of p_T in $|y|$ bins for prompt $D^{*\pm}$ at 7 TeV. The last bin correspond to the overflow bin. This measurement is performed at 7 TeV and the figure is taken from Reference [26].

- Differential cross sections for prompt $D^{*\pm}$ meson production at 13 TeV by CMS [28] including a comparison with the results of ATLAS [25] are shown on the left in

Fig. 5.9a. Here, the data from CMS and ATLAS are compared with FONLL predictions. The two ratios show the individual differences between the theory predictions and the CMS or ATLAS data, respectively.

- These differential cross sections for prompt $D^{*\pm}$ meson production at 13 TeV by CMS [28] including a comparison with the theory prediction by FONLL are shown in Fig. 5.9b. Here, the results are compared with FONLL and several MC simulation models. The measurement is performed differentially in p_T and in the rapidity range up to $|\eta|(D^{*\pm}) < 2.1$. The binning is equidistant with a bin width of 1 GeV up to 8 GeV. This measurement is also performed in 10 equidistant bins of η , which can be found in Reference [28]. For the extraction of the prompt component in the inclusive D^* cross section measurement, the non-prompt background fraction is estimated by simulations co-checked against data. For this purpose, minimum-bias events were used with the PYTHIA 8 tune CUETP8M1 [29].
- Within CMS, the Heavy Ion (HIN) working group measured the prompt [23] and as well the non-prompt [24] component in D^0 final states. This single differential cross section measurement is performed differentially in p_T in the central rapidity range $|y| < 1$ at 5 TeV. In contrast to HIN, the measurement of this thesis is performed with D^* final states, but both results from HIN are used for a comparison with the results of this thesis. The method is explained in more detail in Chapter 6.

ALICE

- The single differential cross section measurement of prompt D^* production from ALICE [27] is shown in Fig. 5.11. It is performed at 7 TeV in the central rapidity range ($|y| < 0.5$). A comparison with the CMS results of Reference [26] is shown in Fig. 5.8. The calculation of the correction for non-prompt D^* production is here performed with a FONLL-based method [27].
- A measurement at 5 TeV for the production prompt D^* is performed by [21] in the central rapidity range $|y| < 0.5$ and p_T range $1 < 36$ GeV. Its results are used for a comparison presented in Chapter 6.
- Differential cross sections for prompt and non-prompt D^0, D^+, D_s^+ are measured at 5 TeV by ALICE [30] in the central rapidity range $|y| < 0.5$ and in p_T up to 36 GeV. For prompt D^0 and prompt D^+ a measurement is even obtained starting at 0 GeV. The

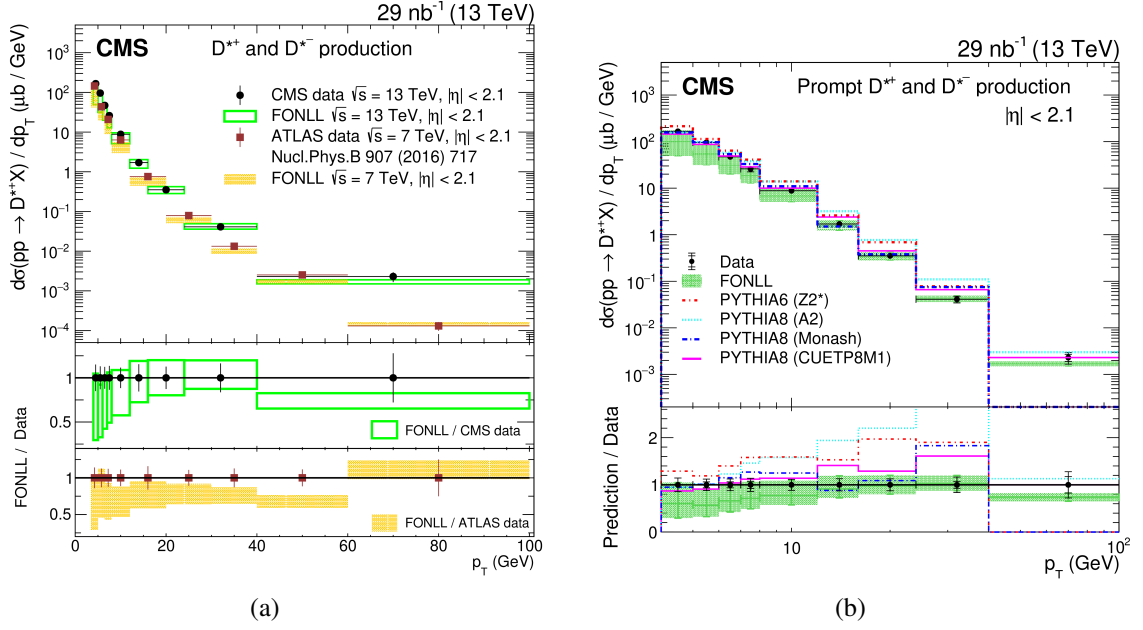


Fig. 5.9 In a), comparison of differential cross sections for $D^{*\pm}$ meson production from CMS (black circles, prompt [28]) at $\sqrt{s} = 13$ TeV with ATLAS (red squares, prompt + non-prompt) at $\sqrt{s} = 7$ TeV [25]. The corresponding predictions from FONLL are shown by the unfilled and filled boxes, respectively. The vertical lines on the points give the total uncertainties in the data, and the horizontal lines show the bin widths. The two lower panels in each plot give the ratios of the FONLL predictions to the CMS and ATLAS data, shown by circles and squares, respectively.

In b), here the non-prompt D^* component is subtracted and differential cross sections for prompt $D^{*\pm}$ meson production at $\sqrt{s} = 13$ TeV are shown. Black markers represent the data and are compared with several MC simulation models and theoretical predictions. The statistical and total uncertainties are shown by the inner and outer vertical lines, respectively. The FONLL band represents the standard uncertainties in the prediction. The lower panel gives the ratios of the predictions to the data. The figures including their captions, which are modified here, are taken from CMS [28].

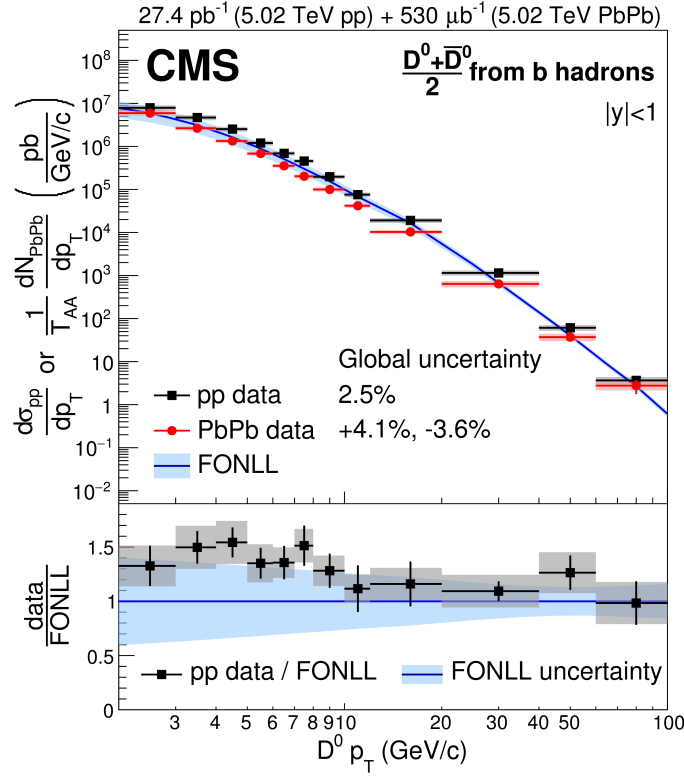


Fig. 5.10 The upper panel shows the differential cross section in pp collisions and the invariant yield in PbPb collisions normalized with T_{AA} at $\sqrt{s} = 5.02$ TeV for $B \rightarrow D^0$ from CMS [24]. The vertical bands around the data points represent the bin-by-bin systematic uncertainties. Uncertainties are smaller than the symbols in most cases. The cross section in pp collisions is compared to FONLL calculations [12]. The lower panel shows the data/FONLL ratio for the $B \rightarrow D^0$ p_T spectra in pp collisions. This figure is taken from CMS [24] and its description modified.

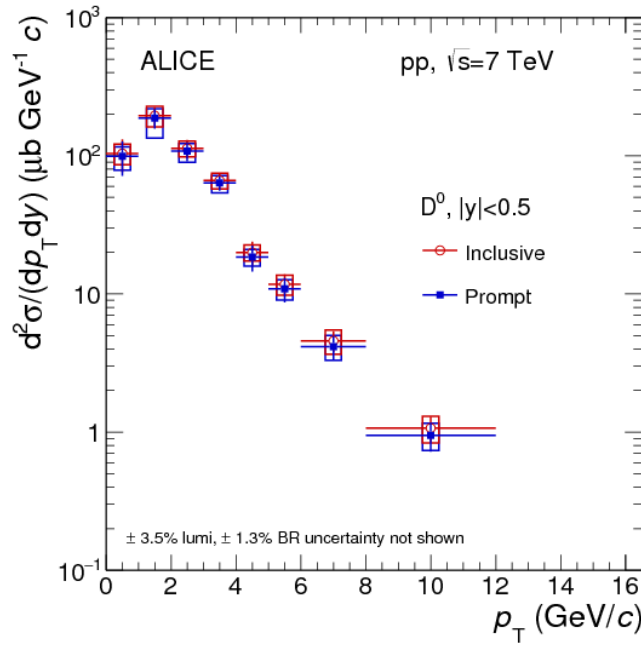


Fig. 5.11 p_T -differential production cross section of D^0 mesons with $|y| < 0.5$ in pp collisions at $\sqrt{s} = 7$ TeV. Comparison of prompt and inclusive D^0 mesons (the latter including also D^0 mesons from beauty-hadron decays) from the analysis without decay-vertex reconstruction. Here the symbols are plotted at the centre of the p_T intervals (shown by the horizontal lines), the vertical lines represent the statistical uncertainties and the vertical size of the boxes corresponds to the systematic uncertainties. This figure and its modified description is taken from ALICE [27].

results are shown in Fig. 5.12. The non-prompt component is here calculated based on a machine-learning approach using a novel-data driven approach [30].

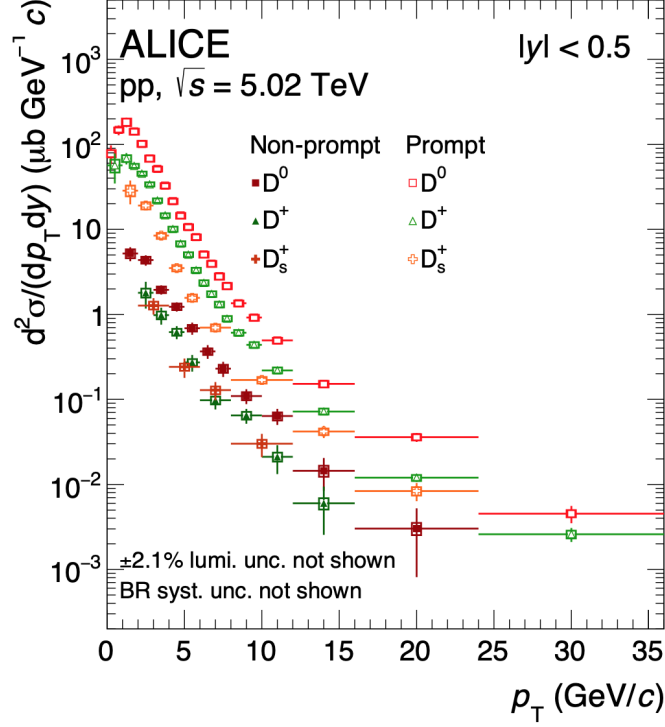


Fig. 5.12 Differential production cross sections for D mesons in the central rapidity range $|y| < 0.5$ in pp collisions at 5 TeV from ALICE [30]. The measurement is performed for prompt and non-prompt D^0, D^+, D_s^+ in p_T up to 36 GeV. This figure is taken from ALICE [30].

LHCb

The double differential cross section measurement by LHCb for the production of prompt D^* is shown in Fig. 5.13. It is performed at 5 TeV and covers the forward region in rapidity $2 < |y| < 4.5$. The extraction of the prompt component in D^* final states is performed via a fit procedure containing a template for the combinatorial background which is obtained by D^* candidates outside the signal region in this analysis [22].

As shown in this chapter, the non-prompt component in D^* final states is only measured at 5 TeV by HIN in the central rapidity range. The next chapter shows the measurement of prompt and non-prompt D^* production in rapidity up to $|y| < 2$.

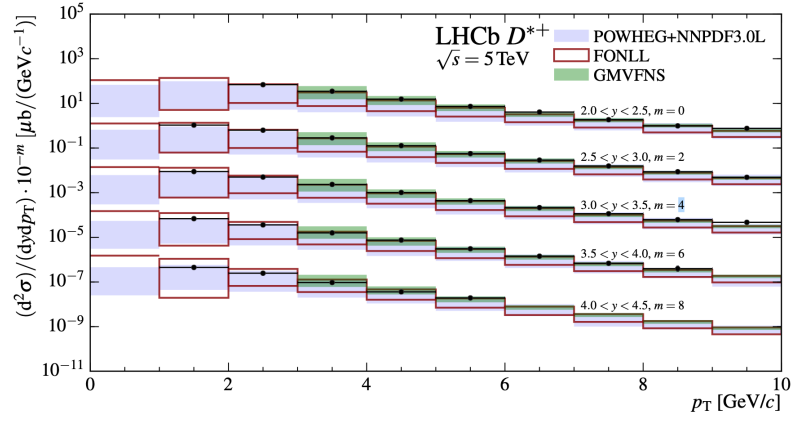


Fig. 5.13 Measurements and predictions for prompt $D^{*\pm}$ cross sections at $\sqrt{s} = 5$ TeV. Each set of measurements and predictions in a given rapidity bin is offset by a multiplicative factor 10^m , where the factor m is shown on the plot. The boxes indicate the $1 \pm \sigma$ uncertainty band on the theory predictions, where only the upper edge is shown if the uncertainty exceeds two orders of magnitude. The figure including its description, which is partially modified, is taken from LHCb [22].

References

- [1] L. Gladilin, “Fragmentation fractions of c and b quarks into charmed hadrons at LEP”, *The European Physical Journal C* **75** (2015), no. 1, 19, 10.1140/epjc/s10052-014-3250-3.
- [2] Fermilab News, “Acronyms of High-Energy Physics”.
<https://www.fnal.gov/pub/ferminews/ahep.html>.
- [3] Belle Collaboration, “Measurement of the Branching Fraction $\mathcal{B}(\Lambda_c^+ \rightarrow pK^-\pi^+)$ ”, *Phys. Rev. Lett.* **113** (2014), no. 4, 042002, 10.1103/PhysRevLett.113.042002, arXiv:1312.7826.
- [4] CLEO Collaboration, “Measurement of $\mathcal{B}(\Lambda_c^+ \rightarrow pK^-\pi^+)$ ”, *Phys. Rev. D* **62** (2000) 072005, 10.1103/PhysRevD.62.072005, arXiv:hep-ex/0004001.
- [5] P. D. Group, “Review of Particle Physics”, *Progress of Theoretical and Experimental Physics* **2020** (08, 2020) 10.1093/ptep/ptaa104. 083C01.
- [6] T. Sjostrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual”, *JHEP* **05** (2006) 026, 10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.
- [7] T. Sjöstrand, S. Mrenna, and P. Skands, “A brief introduction to PYTHIA 8.1”, *Computer Physics Communications* **178** (2008), no. 11, 852 – 867, <https://doi.org/10.1016/j.cpc.2008.01.036>.
- [8] M. G. Bowler, “e+ e- Production of Heavy Quarks in the String Model”, *Z. Phys.* **C11** (1981) 169, 10.1007/BF01574001.
- [9] M. Cacciari et al., “Theoretical predictions for charm and bottom production at the LHC”, *JHEP* **10** (2012) 137, 10.1007/JHEP10(2012)137, arXiv:1205.6344.
- [10] M. Cacciari, M. L. Mangano, and P. Nason, “Gluon PDF constraints from the ratio of forward heavy-quark production at the LHC at $\sqrt{S} = 7$ and 13 TeV”, *Eur. Phys. J. C* **75** (2015), no. 12, 610, 10.1140/epjc/s10052-015-3814-x, arXiv:1507.06197.
- [11] M. Cacciari, S. Frixione, and P. Nason, “The p(T) spectrum in heavy flavor photoproduction”, *JHEP* **03** (2001) 006, 10.1088/1126-6708/2001/03/006, arXiv:hep-ph/0102134.
- [12] M. Cacciari, M. Greco, and P. Nason, “The P(T) spectrum in heavy flavor hadroproduction”, *JHEP* **05** (1998) 007, 10.1088/1126-6708/1998/05/007, arXiv:hep-ph/9803400.
- [13] M. Cacciari, “FONLL Heavy Quark Production”.
<http://www.lpthe.jussieu.fr/~cacciari/fonll/fonllform.html>.
- [14] M. Cacciari, P. Nason, and C. Oleari, “A Study of heavy flavored meson fragmentation functions in e+ e- annihilation”, *JHEP* **04** (2006) 006, 10.1088/1126-6708/2006/04/006, arXiv:hep-ph/0510032.

- [15] M. Cacciari, P. Nason, and R. Vogt, “QCD predictions for charm and bottom production at RHIC”, *Phys. Rev. Lett.* **95** (2005) 122001, 10.1103/PhysRevLett.95.122001, arXiv:hep-ph/0502203.
- [16] M. Cacciari and P. Nason, “Charm cross-sections for the Tevatron Run II”, *JHEP* **09** (2003) 006, 10.1088/1126-6708/2003/09/006, arXiv:hep-ph/0306212.
- [17] P. M. Nadolsky et al., “Implications of CTEQ global analysis for collider observables”, *Phys. Rev. D* **78** (2008) 013004, 10.1103/PhysRevD.78.013004, arXiv:0802.0007.
- [18] V. G. Kartvelishvili, A. K. Likhoded, and V. A. Petrov, “On the Fragmentation Functions of Heavy Quarks Into Hadrons”, *Phys. Lett. B* **78** (1978) 615–617, 10.1016/0370-2693(78)90653-6.
- [19] Yewon Yang, Achim Geiser, Josry Metwally, Valentina Mariani, “Comparison of prompt and non-prompt inclusive charm production cross sections- Yewon Yang, Achim Geiser, Josry Metwally, Valentina Mariani - Production and Properties meeting”, https://indico.cern.ch/event/1026278/contributions/4309065/attachments/2224512/3767461/d_meson_comparison_Yang.pdf.
- [20] Inspire HEP, “Discover High-Energy Physics Content”. <https://inspirehep.net>.
- [21] ALICE Collaboration, “Measurement of D^0 , D^+ , D^{*+} and D_s^+ production in pp collisions at $\sqrt{s} = 5.02$ TeV with ALICE”, *Eur. Phys. J. C* **79** (2019), no. 5, 388, 10.1140/epjc/s10052-019-6873-6, arXiv:1901.07979.
- [22] LHCb Collaboration, “Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 5$ TeV”, *JHEP* **06** (2017) 147, 10.1007/JHEP06(2017)147, arXiv:1610.02230.
- [23] CMS Collaboration, “Nuclear modification factor of D^0 mesons in PbPb collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, *Phys. Lett. B* **782** (2018) 474–496, 10.1016/j.physletb.2018.05.074, arXiv:1708.04962.
- [24] CMS Collaboration, “Studies of Beauty Suppression via Nonprompt D^0 Mesons in Pb-Pb Collisions at $Q^2 = 4$ GeV²”, *Phys. Rev. Lett.* **123** (2019), no. 2, 022001, 10.1103/PhysRevLett.123.022001, arXiv:1810.11102.
- [25] ATLAS Collaboration, “Measurement of $D^{*\pm}$, $D^\pm$ and $D_s^\pm$ meson production cross sections in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector”, *Nucl. Phys. B* **907** (2016) 717–763, 10.1016/j.nuclphysb.2016.04.032, arXiv:1512.02913.
- [26] N. Z. Jomhari, “Measurement of Total and Differential Charm Cross Sections at 7 TeV with the CMS Detector”, 2022. Presented 2022-03-16. <https://cds.cern.ch/record/2805338>.
- [27] ALICE Collaboration, “ D -meson production in p -Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV and in pp collisions at $\sqrt{s} = 7$ TeV”, *Phys. Rev. C* **94** (2016), no. 5, 054908, 10.1103/PhysRevC.94.054908, arXiv:1605.07569.

-
- [28] CMS Collaboration, “Measurement of prompt open-charm production cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *JHEP* **11** (2021) 225, 10.1007/JHEP11(2021)225, arXiv:2107.01476.
- [29] CMS Collaboration, “Event generator tunes obtained from underlying event and multiparton scattering measurements”, *Eur. Phys. J. C* **76** (2016), no. 3, 155, 10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.
- [30] ALICE Collaboration, “Measurement of beauty and charm production in pp collisions at $\sqrt{s} = 5.02$ TeV via non-prompt and prompt D mesons”, *JHEP* **05** (2021) 220, 10.1007/JHEP05(2021)220, arXiv:2102.13601.

Chapter 6

Charm and Beauty Cross Sections

Contents

6.1	Data and Monte Carlo Samples	86
6.1.1	Effective integrated luminosity for pp minimum bias data at 5 TeV	88
6.1.2	Effective luminosity for MC samples	88
6.2	Event Selection and D^* Reconstruction	89
6.3	Background Estimation and Signal Extraction	94
6.4	Determination of the Efficiency	96
6.5	Double Differential Cross Sections for Inclusive D^* production	99
6.6	Separation of Charm and Beauty in D Meson Final States	103
6.7	A Different Ansatz for the Charm Beauty Separation	112
6.8	Double Differential Prompt and Non-prompt D^* Cross Sections	114
6.9	Systematic Uncertainties	122
	References	123

This chapter shows the details for the determination of the cross sections for inclusive *charm* and *beauty* production at the LHC at 5 TeV down to very low transverse momentum, and the comparison with QCD predictions in next-to-leading order of perturbation theory. The measurement of the cross sections for the production of heavy quarks at the LHC is one crucial test of QCD, and can for instance, as has already happened in the case of top production [1], be used for a measurement of the quark masses. As shown in Chapter 5 other experiments like ATLAS and ALICE covered only small fractions of the available phase space with central rapidity while the LHCb experiment covered the forward region, $2.0 < y < 4.5$.

In this work, cross sections are measured double differentially in almost the full phase space complementary to LHCb, of prompt D^* mesons, and D^* mesons from b hadron decays through the decays $B \rightarrow D^* X$. One of the challenges is the separation of prompt D^* mesons and D^* mesons from b hadron decays near the production threshold, called charm-beauty separation in this thesis. This separation is a necessary ingredient for the measurement of cross sections for prompt and non-prompt D^* , eventually leading to the determination of the total cross section for inclusive *charm* and *beauty* production, respectively.

This chapter starts with an overview of the data used and Monte Carlo samples entering this work and continues with a list of explicit cuts used for this analysis. The procedure of the event selection and of the reconstruction of D^* candidates, which is performed in small bins in p_T and $|y|$ (rapidity) in the full accessible phase space from CMS, can be found in Section 6.2. The signal extraction and background estimation can be found in Section 6.3 and the determination of the efficiency for the reconstruction of prompt and non-prompt D^* in Section 6.4. The resulting double differential cross sections for inclusive D^* are shown in Section 6.5 for $0 < |y| < 2.5$. The procedure for the charm-beauty separation and its application is illustrated and documented in Section 6.6. The resulting double differential cross sections for prompt and non-prompt D^* mesons are shown in Section 6.8. A discussion about the considered systematic uncertainties related to this measurement can be found in Section 6.9. Through the combination of the measurements of LHCb with those from this work, and filling gaps of the phase space with theory predictions, total cross sections can be extracted.

6.1 Data and Monte Carlo Samples

This section lists details about the Run-II data and Monte Carlo samples and the technical extraction of their effective luminosity used in this study. For the measured cross sections within this thesis, minimum bias data (MB data) recorded in Run-II at 5 TeV from 2015 are primarily analyzed. MB data are events recorded with as little bias as possible c.f. Section 6.1.1. The number of processed events of the used MB data, including the integrated luminosity, are shown in Tab. 6.1. The integrated luminosity is shown in the last row and the last column represents the number of minimum bias collisions. The table shows the names of 2015E 5 TeV pp datasets and the number of events without and with using a JSON file. A JSON file lists the selection of good events i.e. for those events the detector configuration worked under good conditions. The further content and use of a JSON file is explained in Section 6.1.1. The MC samples generated with PYTHIA6 produced for this work are listed

in Tab. 6.3. For the measurement of charm and beauty cross sections, the $D^0 \rightarrow K\pi$ MC sample is used in this study.

Table 6.1 This table lists the different data samples used in this work. The list of 2015E 5 TeV pp datasets shows the name of the dedicated sample, the number of events without using a JSON file, and in the next column the next number with using it. The last column represents the minimum bias collisions, i.e. here the number of reconstructed good quality primary vertices (Section 6.2). The integrated luminosity is shown in the last row.

Sample	#Events	w. JSON	MB
/MinimumBias1/Run2015E-PromptReco-v1/AOD	126,809,757	111,413,194	
/MinimumBias2/Run2015E-PromptReco-v1/AOD	126,998,875	111,600,855	
/MinimumBias3/Run2015E-PromptReco-v1/AOD	126,853,017	111,502,472	
/MinimumBias4/Run2015E-PromptReco-v1/AOD	127,250,025	111,851,720	
/MinimumBias5/Run2015E-PromptReco-v1/AOD	127,169,537	111,772,459	
/MinimumBias6/Run2015E-PromptReco-v1/AOD	127,256,729	111,859,378	
/MinimumBias7/Run2015E-PromptReco-v1/AOD	127,256,692	111,858,963	
/MinimumBias8/Run2015E-PromptReco-v1/AOD	127,239,988	111,841,797	
/MinimumBias9/Run2015E-PromptReco-v1/AOD	127,222,974	111,826,347	
/MinimumBias10/Run2015E-PromptReco-v1/AOD	127,220,628	111,822,249	
/MinimumBias11/Run2015E-PromptReco-v1/AOD	126,325,160	111,857,169	
/MinimumBias12/Run2015E-PromptReco-v1/AOD	127,207,059	111,808,958	
/MinimumBias13/Run2015E-PromptReco-v1/AOD	125,206,184	109,958,587	
/MinimumBias14/Run2015E-PromptReco-v1/AOD	126,522,737	111,602,374	
/MinimumBias15/Run2015E-PromptReco-v1/AOD	126,753,153	111,602,223	
/MinimumBias16/Run2015E-PromptReco-v1/AOD	127,128,323	111,872,629	
/MinimumBias17/Run2015E-PromptReco-v1/AOD	126,280,043	111,105,435	
/MinimumBias18/Run2015E-PromptReco-v1/AOD	126,542,929	111,311,694	
/MinimumBias19/Run2015E-PromptReco-v1/AOD	126,373,548	111,214,484	
/MinimumBias20/Run2015E-PromptReco-v1/AOD	127,031,373	111,633,897	
Total	2,536,648,731	2,231,316,884	1,853,304,000
effective luminosity (see Section 6.1.1)		40.17 nb ⁻¹	40.17 nb ⁻¹

Table 6.2 This table lists the JSON files of 2015 5 TeV proton proton collision. The symbol (*) represents the one used by CMS-HIN [2, 3] (Heavy Ion working group within CMS with which results are compared within this study) while the symbol (**) stands for the latest JSON file that is used within this work leading to more statistics compared to (*).

JSON file
*Cert_262081-262273_5TeV_PromptReco_Collisions15_25ns_JSON_v2.txt
**Cert_262081-262328_5TeV_PromptReco_Collisions15_25ns_JSON.txt

6.1.1 Effective integrated luminosity for pp minimum bias data at 5 TeV

In pp collisions, the cross section for charm production is predicted to be $\approx 10 \text{ mb}^{-1}$ and for beauty production $\approx 1 \text{ mb}^{-1}$ [4]. On average, in each Minimum Bias event there is ≈ 1 pp collision (very little pileup), but not every event has collisions leading to a reconstructable primary vertex. To increase the statistics, also pileup vertices are used in this analysis, since they can contain charm and beauty events. This happened on average about in 10% of all used events. The determination of the integrated luminosity for the 2015 5 TeV pp data of 40.17 nb^{-1} used in this work is obtained using Brilcalc [5]. The Minimum bias datasets shown in Tab. 6.1 were recorded via the high level trigger (HLT)

HLT_L1MinimumBiasHF10R_part*_v1

where * stands for 0 to 19. Here, events are recorded when an activity is noticed by the HF of the HCAL shown in Fig. 3.4. In general an event can have several collisions and huge pileup. In this thesis, a minimum bias event has almost no pileup. The used data is based on only this one active trigger to record each event. Prescale factors which are used to reduce the trigger rate are taken into account by Brilcalc to get the correct luminosity for each individual sample shown in Tab. 6.1. Brilcalc is taking the JSON file shown in Tab. 6.2 to determine the effective integrated luminosity.

In Appendix B.1 the executed command with its entire output is shown. The final luminosity and the uncertainty on the luminosity value is obtained to be 40.17 nb^{-1} and 2.3 %, respectively, based on the public note LUM-16-001 [6]. More information about this luminosity measurement can be found in Reference [6].

6.1.2 Effective luminosity for MC samples

All the information that is shown in this section and even more can be found on the general PdmV¹ web interface [7] for McM². The effective luminosity for the MC samples has to be calculated individually via

$$\mathcal{L}_{eff} = \frac{N}{\sigma}, \quad (6.1)$$

where σ is the cross section corresponding to the physical process with the amount of the signal N for this process. The PdmV web interface [8] contains information about the produced MC samples by CMS. For instance the underlying cross section of the 2015 5 TeV

¹Physics Data and Monte Carlo Validation

²Monte Carlo Manager; Monte Carlo management tool

$D^0 \rightarrow K\pi$ MC sample shown in Tab. 6.3 is

$$\sigma_{pp \rightarrow D^0 \rightarrow K\pi} = 298,7 \mu\text{b}, \quad (6.2)$$

which leads to an effective luminosity of

$$\mathcal{L}_{eff} = \frac{N}{\sigma} = \frac{12077624}{298,7 \mu\text{b}} = 40,43 \text{ nb}^{-1}. \quad (6.3)$$

The value for this effective luminosity of the $D^0 \rightarrow K\pi$ sample can be found in Tab. 6.3. This sample is designed to have no pileup and at least one D^0 decaying into a K and π . Other MC samples are listed in this table too. They are needed for instance for a systematic study of the background (Section 6.9). Here, quark filters were used for the individual generated MC samples. A quark filter is used to make sure to have at least one dedicated heavy quark (here charm or beauty) in the production process. For the *MinBias charmfilter* and *MinBias beautyfilter* MC sample, the charm filter does not exclude additional beauty, and vice versa. The D^0 and as well the D^+ samples contain both charm and beauty (prompt and non-prompt).

Table 6.3 List of 2015 5 TeV pp MC generated through CMS-BPH (B Physics working group within CMS), full phase space. The table shows the name of the MC sample in the first column, its number of events in the second and in the last column the effective luminosity.

Sample	#Events	$\mathcal{L}_{eff}$ (Section 6.1.2)
/MinBias_charmfilter_TuneCUEP8M1_5TeV_pythia8-evtgen/ HINppWinter16DR-75X_mcRun2_asymptotic_ppAt5TeV_v3-v3/AODSIM	44,697,355	8.94 nb ⁻¹
/MinBias_beautyfilter_TuneCUEP8M1_5TeV_pythia8-evtgen/ HINppWinter16DR-75X_mcRun2_asymptotic_ppAt5TeV_v3-v3/AODSIM	4,810,667	18.7 nb ⁻¹
/DOKpi_pT0toInf_TuneCUEP8M1_5TeV_pythia8-evtgen/ HINppWinter16DR-75X_mcRun2_asymptotic_ppAt5TeV_v3-v3/AODSIM	12,077,624	40.4 nb ⁻¹
/DplusKpipi_pT0toInf_TuneCUEP8M1_5TeV_pythia8-evtgen/ HINppWinter16DR-75X_mcRun2_asymptotic_ppAt5TeV_v3-v3/AODSIM	11,994,686	36.8 nb ⁻¹

6.2 Event Selection and D^* Reconstruction

In this analysis, each D^* is reconstructed via its decay products $D^{*\pm} \rightarrow D^0 \pi_s^\pm \rightarrow K^\mp \pi^\pm \pi_s^\pm$. The D^* is decaying into D^0 and π_s , where the index s stands for slow or soft, since its momentum is very small due to the tiny phase space remaining from the very low mass difference of the D^0 and D^* shown in Tab. 5.2. Subsequently, the D^0 is decaying into K and a π , where the K carries always the opposite charge compared to π_s . The exact number of used

collisions corresponds to the number of vertices satisfying the criteria for the reconstructed primary vertex. In the following the main criteria are listed:

- take a vertex from the primary vertex collection with beam spot constraint
- this vertex has to be valid i.e. the fit converged reasonably
- not fake: at least one track
- the absolute distance in the z-direction between the primary vertex and the beam spot has to be: $|d_{z, \text{vertex}} - d_{z, \text{beam spot}}| < 15 \text{ cm}$
- and the distance in the xy-direction between the vertex and beam spot:
 $|(d_{xy})d_{T, \text{vertex}} - d_{T, \text{beam spot}}| < 2 \text{ cm}$

In this analysis in the case of prompt D^* mesons, the primary vertex corresponds to the D^* vertex and the decay vertex of the resulting D^0 to the secondary vertex (SV). This is illustrated in Fig. 6.1. The SV is fitted by using `KalmanFitterVertex` [9] based on the Kalman-filter algorithm explained in Section 4.1. It takes two tracks with an opposite charge, which correspond to the K and π . If the fitting procedure of the `KalmanFitterVertex` converges, the position of the vertex and its uncertainty are determined. Together with the refitted tracks, the D^0 candidate and the D^0 vertex determine the SV in this analysis. For the reconstruction of a prompt D^* candidate, the D^0 needs to be associated with an additional track corresponding to the π_s . It originates from the corresponding PV, which is the same vertex where the D^0 candidate originates from, while for a non-prompt D^* it does not come from the PV. A number of additional variables are used in reconstruction and selection of D^* candidates, which are defined in the following.

The fraction of the momenta $p_{T, \text{frac}}$ corresponds to the ratio of the transverse momentum of the D^0 and the sum of all transverse momenta of all tracks with respect to the PV:

$$p_{T, \text{frac}} = \frac{p_T^D}{p_T \text{ of } \sum_{\text{all tracks at respective PV}}}. \quad (6.4)$$

The decay length of the D^0 is given by

$$dl^{D^0} = \frac{\vec{p}^{D^0} \cdot \vec{d}_{\Delta, \text{SV}}}{|\vec{p}^{D^0}|}, \quad (6.5)$$

where $\vec{p}^{D^0}$ is the three momentum vector of the D^0 candidate, $\vec{d}_{\Delta, \text{SV}}$ the distance vector between the SV and the PV positions. Another useful variable is the decay length significance

given by

$$dl_{sig}^{D^0} = \frac{dl^{D^0}}{dl_{err}^{D^0}}, \quad (6.6)$$

where $dl_{err}^{D^0}$ corresponds to the uncertainty of the decay length of the D^0 . The angle $\cos \phi$ between the decay length dl^{D^0} and the distance $d_{\Delta,SV}$ is given by

$$\cos \phi = \frac{dl}{d_{\Delta,SV}}. \quad (6.7)$$

An illustration of the PV and SV together with the decay of a D^* into a D^0 and π_s is shown in Fig. 6.1. By collecting the candidates for the D^* , only a right charge combination

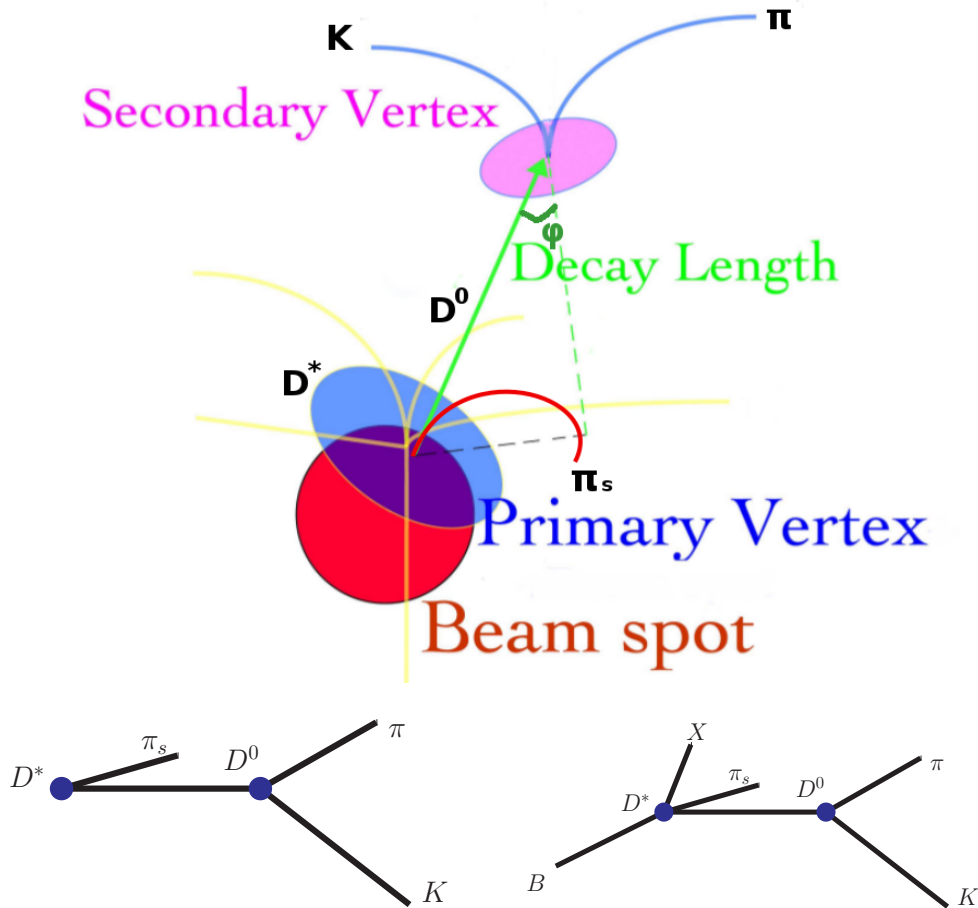


Fig. 6.1 Sketch of the decay $D^* \rightarrow D^0 \pi_s \rightarrow K \pi \pi_s$ and a corresponding cartoon for this decay. The third cartoon shows the D^* coming from a b hadron.

of its decay products (π_s, K, π) can lead to a real D^* . Here, the a charge combination is made

that the kaon and pion have opposite charge as they came from D^0 . It is conventional that a D^0 corresponds to $K^-\pi^+$ and a $\bar{D}^0$ to $K^+\pi^-$. Secondly, the right charge combination means that the charge of the D^* corresponds to the one from the π_s . A wrong charge combination of this D^* system can be constructed by selecting the combinations where the kaon and pion have the same charge leading to " D^{++}, D^{--} ". These do not exist as exclusive decay modes of D mesons i.e. collecting these non physical combinations leads to a combinatorial background which is free of any D or D^* signal. Secondly, the π_s is then required to have the opposite charge of the " D^{++}, D^{--} ", such that the total charge of the D^* system is 1, as for the right charge combinations. This avoids an overall charge bias. Therefore, this wrong charge combination of the D^* system can be used to determine the background in the signal region. In this data driven way, a very clear signal is extracted as shown in Section 6.3.

The preselection cuts for D^* candidates at AOD level are shown in Tab. 6.4. A cut on p_T is imposed to ensure that the p_T of the D^0 candidate contributes significantly to the total p_T of all tracks associated to the vertex. The cuts on the decay products of the D^* candidates at analysis level are shown in Tab. 6.5. Here, for instance, a cut on dE/dx is performed to take the different energy loss of the particles passing the detector components into account [10]. Furthermore, a selection at analysis level of D^* candidates is performed and their criteria are shown in Tab. 6.6. Here, selection criteria are separated into a lower ($p_T < 3.5$ GeV) and higher ($p_T > 3.5$ GeV) p_T region. To reduce the background, some cuts in the lower p_T region are tighter compared to the higher p_T region.

Table 6.4 Preselection of D^* candidates at AOD level.

Variable	Cut
transverse momentum of K and π $p_T^{K,\pi}$	> 0.3 GeV
d_{xy} and d_z distances between K and π track origins	< 0.5 cm
$p_{T, \text{frac}}$ for $p_T^{D^0} < 0.9$ GeV	> 0.15
z distance between D^0 vertex and PV	< 2 cm
$p_T^{D^0}$	> 0.9 GeV
loose m_D^0 cut	$\in (1.5, 2.3)$ GeV
ratio of $p_T^{\pi_s}$ and $p_T^{D^0}$	$\in (0.03, 0.20)$
d_{xy} and d_z distances between π_s and D^0 vertex	< 2 cm
mass difference between D^* and D^0 , Δm	< 0.17 GeV
'cross' cut on D^0 candidates from D^*	$m_{D^0} \in (1.828, 1.908)$ GeV or $\Delta m \in (0.1434, 0.1474)$ GeV

Table 6.5 Cuts at analysis level on the decay products of D^* .

Variable	Cut
p_T^π	$> 0.5 \text{ GeV}$
p_T^K	$> 0.3 \text{ GeV}$
$d_{xy}^{K,\pi}$ from PV	$< 0.15 \text{ cm}$
$d_z^{K,\pi}$ from PV	$< 0.1 / \sin \theta \text{ cm}$
$d_{xy}^{\pi_s}$ from PV	$< 0.3 \text{ cm}$
$d_z^{\pi_s}$ from PV	$< 0.2 / \sin \theta \text{ cm}$
dE/dx of K for $p^K < 1.5 \text{ GeV}$	$\in [0.6/ p^K + 2, 1.0/ p^K + 3.5]$

Table 6.6 List of criteria for the selection of D^* candidates for the lower and higher p_T region.

	lower p_T	higher p_T
$p_T^{D^*}$	$1.5 - 3.5 \text{ GeV}$	$> 3.5 \text{ GeV}$
m_{D^0}	$1.836 < m_{D^0} < 1.89 \text{ GeV}$	$1.85 < m_{D^0} < 1.88 \text{ GeV}$
dl_{sig}	$p_{T,\text{frac}}^{D^0} > 0.1 \text{ \&}$ $\left\{ \left(dl_{\text{sig}}^{D^0} > 1.5 \text{ \& } p_{T,\text{frac}}^{D^*} > 0.15 \right) \right.$ $\text{or } dl_{\text{sig}}^{D^0} > 3$ $\left. \text{or } \left(dl_{\text{sig}}^{D^0} > 2 \text{ \& } \cos(\phi_{D^0}) > 0.995 \right) \right\}$	$(dl_{\text{sig}} > -1 \text{ \& } p_{T,\text{frac}} > 0.15)$ $\text{or } dl_{\text{sig}} > 2$
$\cos \phi$	0.8	0.8
$\Delta m = m_{K\pi\pi_s} - m_{K\pi}$	$0.14440 < \Delta m < 0.14664 \text{ GeV}$	$0.14440 < \Delta m < 0.14664 \text{ GeV}$

A quality check of the D^* reconstruction can be seen in 6.2. It shows a diagonal correlation for p_T , ϕ and y of the reconstructed D^* candidates matched with the generated D^* based on their kinematics. The migration between the reconstructed and generated quantities is small. Thus, the number of extracted signal events can be directly used with a bin-to-bin unfolding technique and low p_T bins are used within this thesis for the determination of the cross sections. Furthermore, it shows that reconstruction of D^* candidates is possible starting at 1.5 GeV.

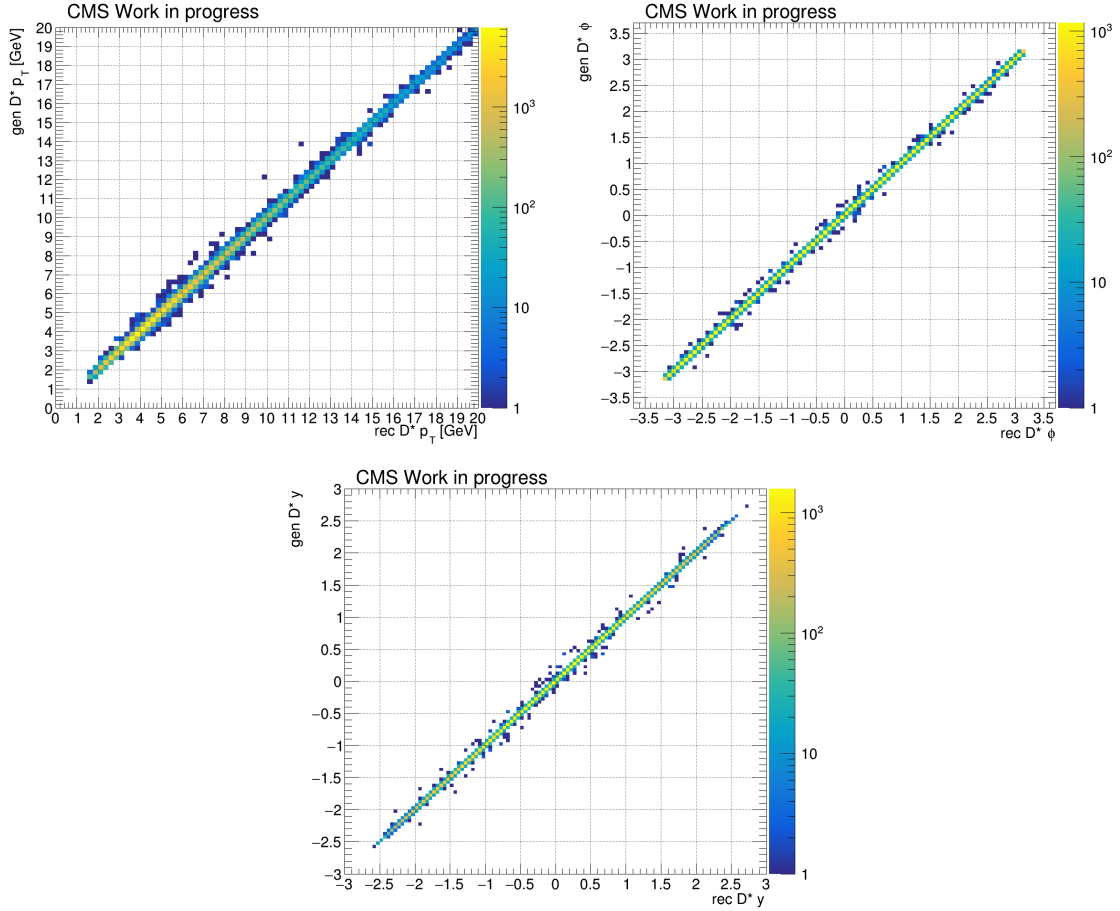


Fig. 6.2 The diagonal correlation for p_T , ϕ and y of the reconstructed D^* candidates matched with the generated D^* based on their kinematics implies that the migration is small.

6.3 Background Estimation and Signal Extraction

Focusing on the Δm distribution ($\Delta m = m(K\pi\pi_s) - m(K\pi_s)$) and combining both, wrong and right charge D^* candidates, the signal extraction of D^* is performed. The right (wrong)

charge combination is shown in blue (red) and the signal (side band) region in pink (gray) in Fig. 6.3. The Δm distributions for the signal extraction for D^* are shown for the lower and higher p_T regions. First, the wrong charge combination distribution (combinatorial background) is normalized to the right charge distribution in the side band region. This leads to a scale factor (SF), which is ≈ 1 in data for combinatorial background. They are listed for all phase space bins in the corresponding Δm distributions. This extracted SF is then applied to the entire wrong charge distribution. In Fig. 6.3 it is already applied. Lastly, the signal is obtained by subtracting the normalized wrong charge combinations from the right charge combinations in the signal region. In the same way, the signal extraction can be performed for the $D^0 \rightarrow K\pi$ MC sample discussed in Section 6.1. The signal extraction for prompt and non-prompt D^* is used in Section 6.6 for the charm beauty separation in D^* final states. Since the used MC is constructed to have per event at least one D^0 , it contains almost no background. The subtraction method leads in this case to a SF which is $\gg 1$.

The application of the subtraction method is illustrated in the higher and lower p_T regions for prompt and non-prompt D^* in the matched MC shown in Fig. 6.6 and Fig. 6.7, respectively.

For illustration purposes, the signal region is also fitted with a function. It contains a threshold function for the background

$$A \cdot (\Delta m - m_\pi)^B \cdot \exp[C \cdot (\Delta m - m_\pi)] \quad (6.8)$$

where A, B and C correspond to free parameters and m_π to the mass of the pion, and the signal distribution is fitted with a modified Gaussian function [11]

$$\text{Gauss}^{\text{mod}} = N \cdot \exp \left[-0.5 \cdot x^{1+1/(1+0.5-x)} \right], \quad (6.9)$$

with $x = |(\Delta m - m_0)/\sigma|$ and N as normalization factor. Here, the signal position m_0 and the width σ correspond to free parameters. The signal extraction in this analysis is however obtained only by the subtraction method based on the right and wrong charge combination. A more precise description of Eq. 6.9 can be found in References [11].

The signal extraction is performed in small bins of p_T and $|y|$ in the full accessible phase space of CMS ($|y| < 2.5$, $p_T > 1.5$ GeV). An overview is shown in Fig. 6.4 and a table with the corresponding numbers including its uncertainty derived by the subtraction method for the signal extraction is shown in Fig. 6.5. The uncertainty of the obtained signal is derived by taking the statistical uncertainties from the right charge and wrong charge into account. A list of all signal extraction plots in data for each bin in p_T and $|y|$ can be found in Appendix B.2. A list of all extraction plots of the signal for prompt and non-prompt production in the used

MC sample for each bin in p_T and $|y|$ can be found in Appendix C.2. With the obtained signal for D^* production, differential cross sections for inclusive D^* production can be calculated, or even more, a separation of prompt and non-prompt D^* leading to their production cross sections can be performed.

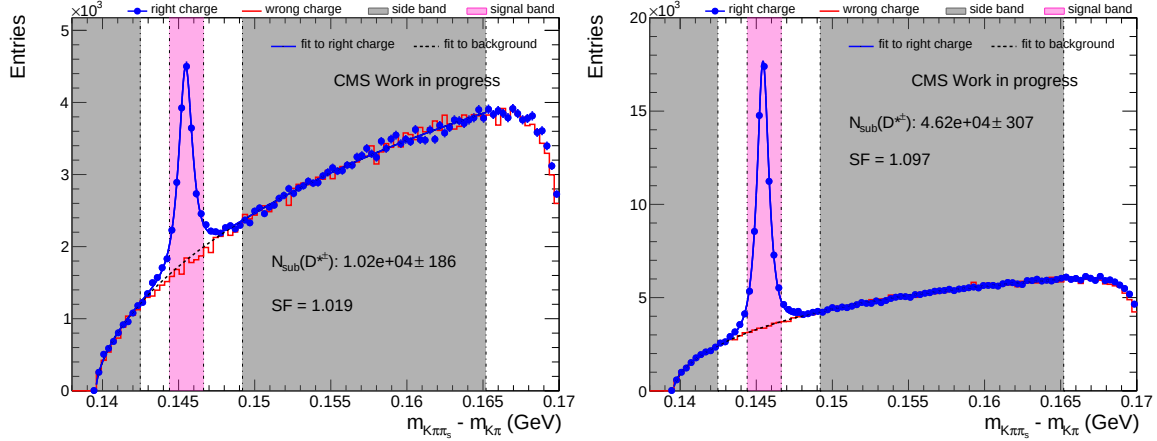


Fig. 6.3 Δm distributions for the signal extraction in the lower (left) and higher (right) p_T region for D^* in the data. Normalizing red to blue in the gray area (side band region) leads to a scale factor (SF) which is then applied to the entire wrong charge combination (red). The signal is then obtained by the subtraction of this product in the signal region (pink) from blue. For illustration purposes, the signal extraction is compared with a fitted modified Gaussian function (see text).

6.4 Determination of the Efficiency

In addition to the signal extraction, the efficiency determination is a necessary ingredient for calculating cross sections. It evaluates the probability with which a D^* is reconstructed in the individual bins of the phase space. For this procedure the $D^0 \rightarrow K\pi$ MC sample shown in Tab. 6.3 is used. The efficiency corresponds to the ratio of reconstructed and true D^* such that

$$\epsilon_{\text{efficiency}} = \frac{N_{\text{reco \& matched}}}{N_{\text{true}}}. \quad (6.10)$$

Here, $N_{\text{reco \& matched}}$ corresponds to the number of reconstructed D^* that are matched with the D^* generated at true level i.e. apart from the matching, cuts are only applied at reconstruction level. This table is for inclusive D^* production shown in Tab. 6.8a and for prompt (non-prompt) D^* production in Tab. 6.9a (Tab. 6.10a). The total number of generated true D^* corresponds to N_{true} . This table is for inclusive D^* production shown in Tab. 6.8b and for

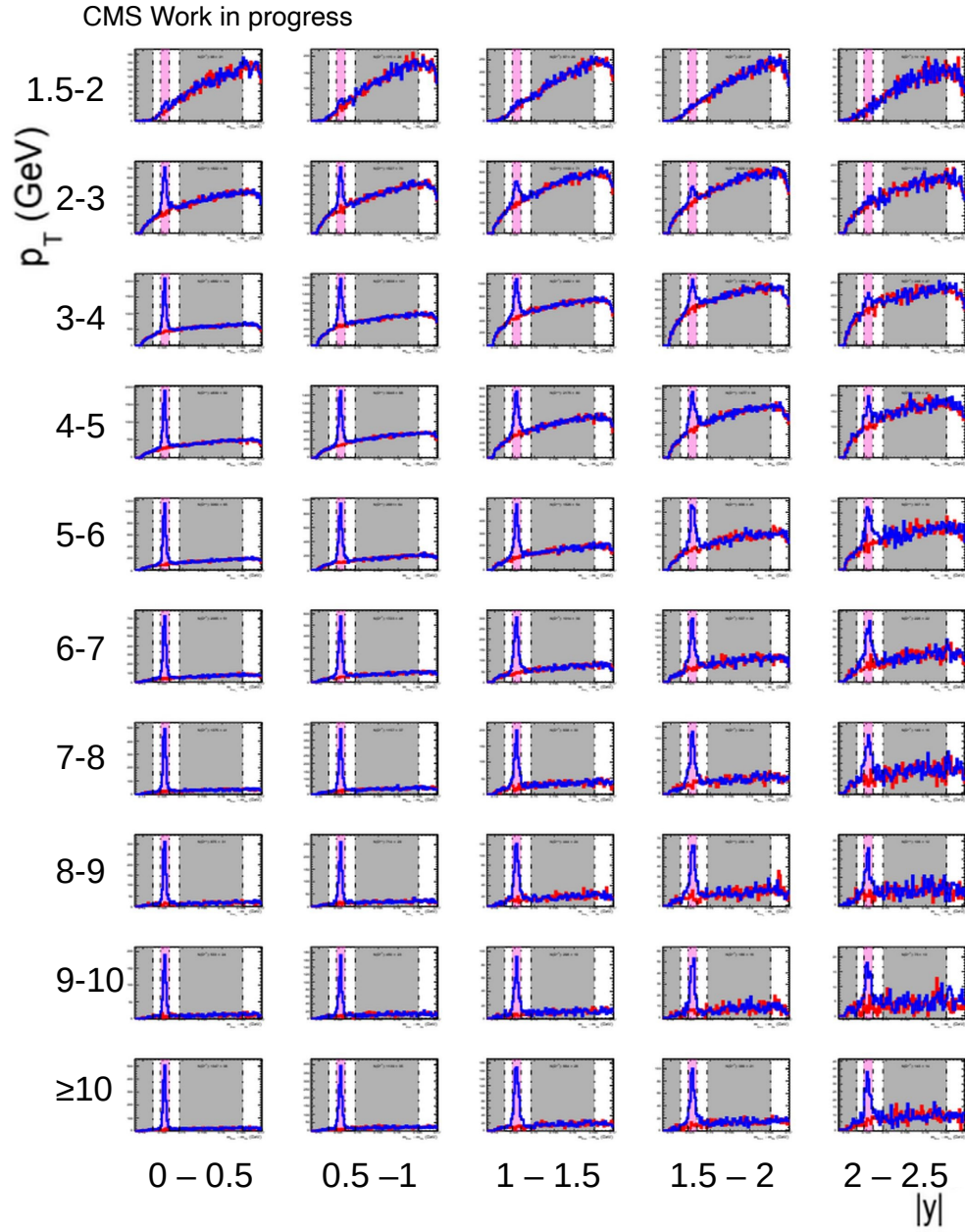


Fig. 6.4 Overview of signal extraction for $D^* \rightarrow K\pi\pi_s$ in small bins in p_T for $|y| < 2.5$ and $p_T > 1.5$ GeV. For enlarged figures see Appendix B.2.

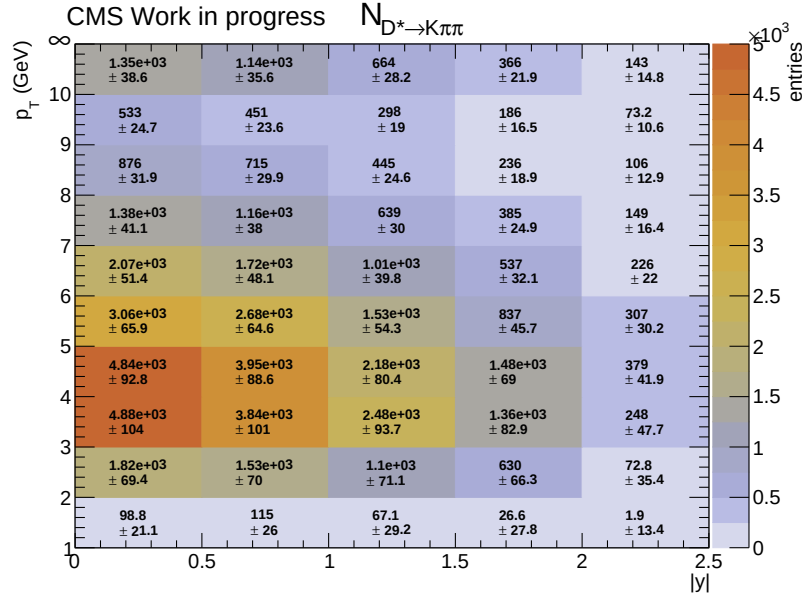


Fig. 6.5 Signal number of events and its statistical uncertainty derived by the subtraction method for $D^* \rightarrow K\pi\pi_s$ in bins in p_T and $|y|$ for $|y| < 2.5$ and $p_T > 1.5$ GeV in data.

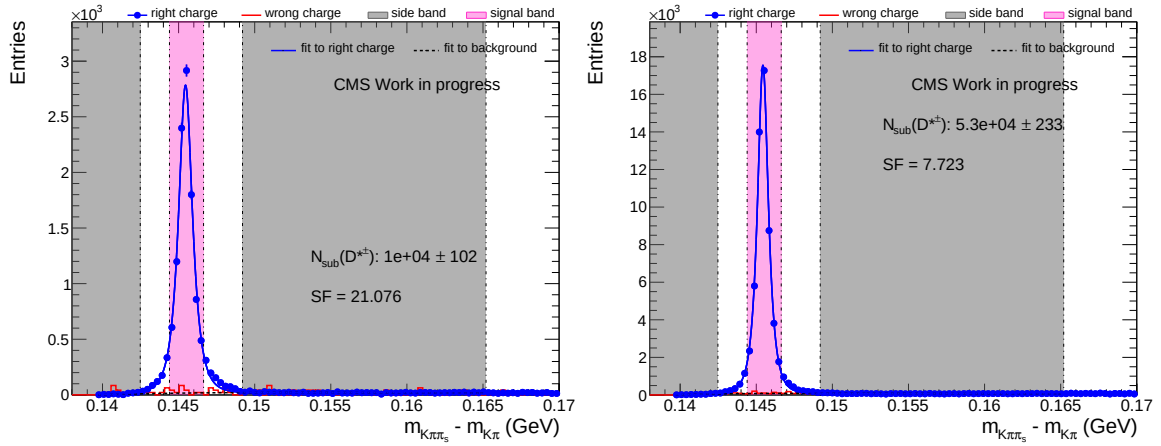


Fig. 6.6 Signal extraction in lower (left) and higher (right) p_T region for prompt D^* in the matched MC. Normalizing red to the blue (in the gray area) leads to a scale factor (SF) which is applied to red (wrong charge combination) and subtracted then in the signal region (pink) from blue (right charge combination) to get the signal. For illustration purposes, a modified Gaussian function is fitted in the signal region, but in this analysis, the signal $N_{sub}(D^{*\pm})$ is extracted via the subtraction method.

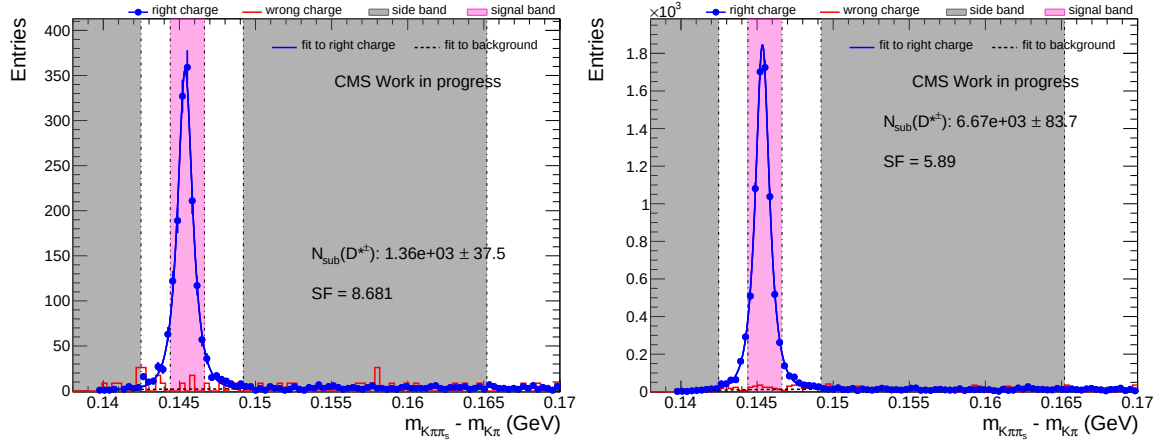


Fig. 6.7 Signal extraction in lower (left) and higher (right) p_T region for non-prompt D^* in the matched MC. Normalizing red to the blue (in the gray area) leads to a scale factor (SF) which is applied to red (wrong charge combination) and subtracted then in the signal region (pink) from blue (right charge combination) to get the signal. For illustration purposes, a modified Gaussian function is fitted in the signal region, but in this analysis, the signal $N_{sub}(D^{*\pm})$ is extracted via the subtraction method.

prompt (non-prompt) D^* production in Tab. 6.9b (Tab. 6.10b). The number of $N_{reco \& matched}$ and N_{true} bins of p_T and $|y|$ can be seen in Fig. 6.8a and Fig. 6.8a, respectively. The detector efficiency table for $D^* \rightarrow K\pi\pi_s$ is shown as a function of p_T and $|y|$ in Fig. 6.8c and its three dimensional distribution in Fig. 6.8d. It shows that for higher p_T in the central rapidity range an efficiency up to 54% is reached, while the efficiency drops at low p_T and for high $|y|$. Since this study aims not only to the calculation of D^* cross sections but of the prompt and non-prompt $D^* \rightarrow K\pi\pi_s$ cross section too, the efficiencies for prompt and non-prompt $D^* \rightarrow K\pi\pi_s$ were calculated accordingly with Eq. 6.10. The result is shown in Fig. 6.9 for prompt D^* and in Fig. 6.10 for non-prompt D^* .

Due to the b hadron lifetime, non-prompt D^* can be better reconstructed compared to the prompt D^* at lower p_T , while for higher p_T this changes. This can be seen in the central rapidity range up to 1 in $|y|$ by comparing Fig. 6.9c with Fig. 6.10c or the three dimensional illustration in Fig. 6.9d with Fig. 6.10d, respectively.

6.5 Double Differential Cross Sections for Inclusive D^* production

This section shows the measurements of inclusive D^* production without the separation of prompt from non-prompt D^* . This measurement is performed in the phase space $p_T > 1.5$

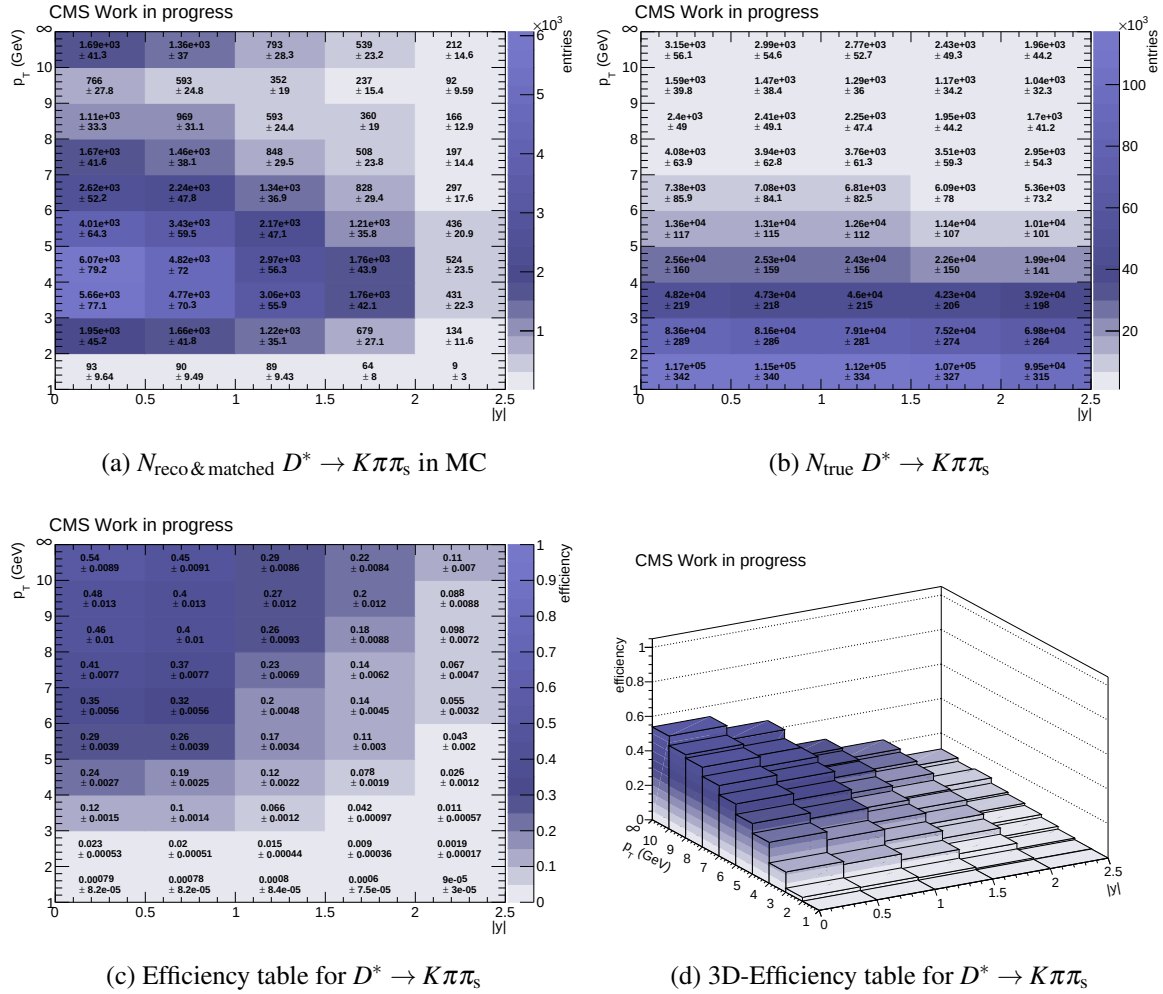


Fig. 6.8 Efficiency table and its statistical uncertainty for $D^* \rightarrow K\pi\pi_s$ reconstruction in bins of p_T and y for $|y| < 2.5$, $p_T > 1.5$ GeV.

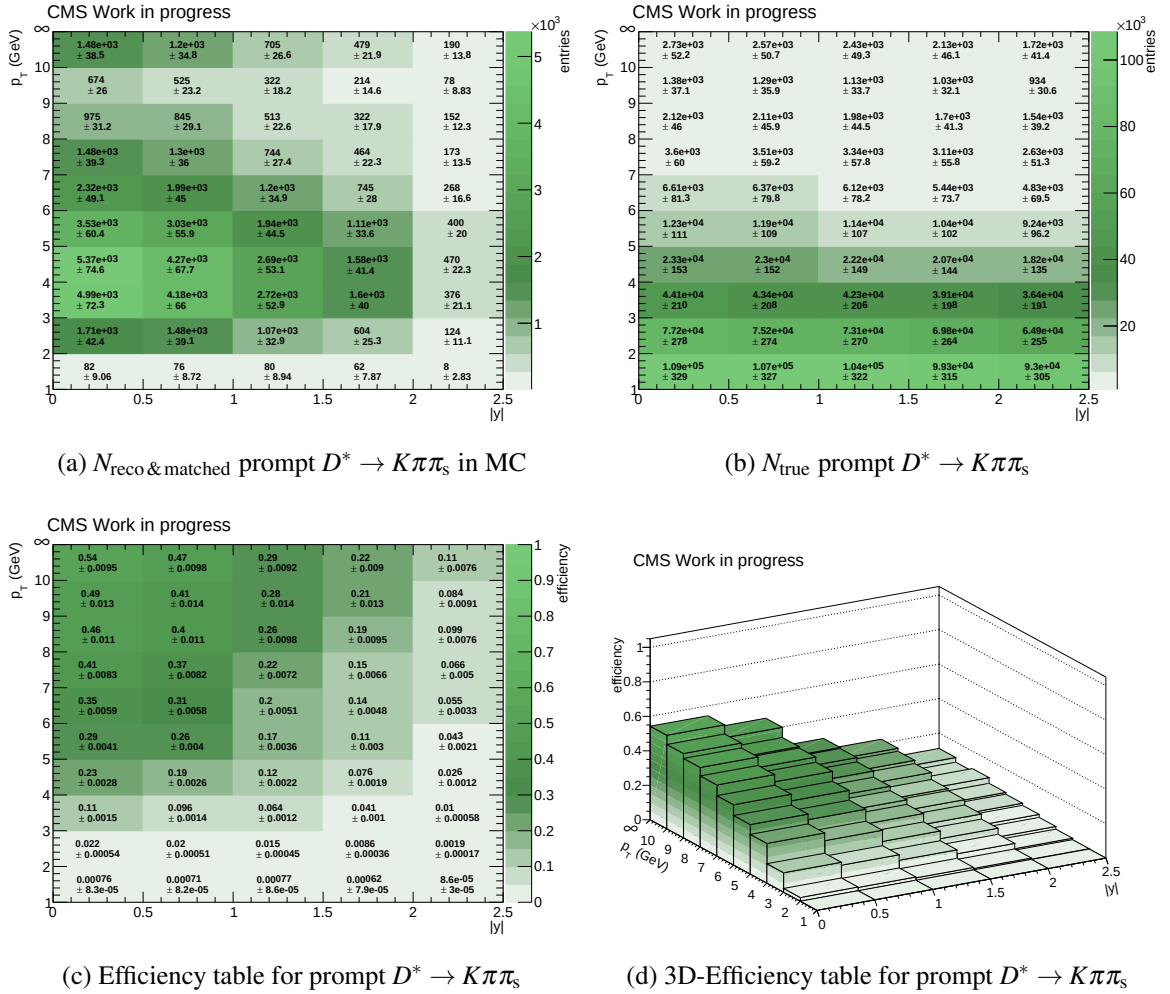


Fig. 6.9 Efficiency table and its statistical uncertainty for prompt $D^* \rightarrow K\pi\pi_s$ reconstruction in bins of p_T and y for $|y| < 2.5$, $p_T > 1.5$ GeV.

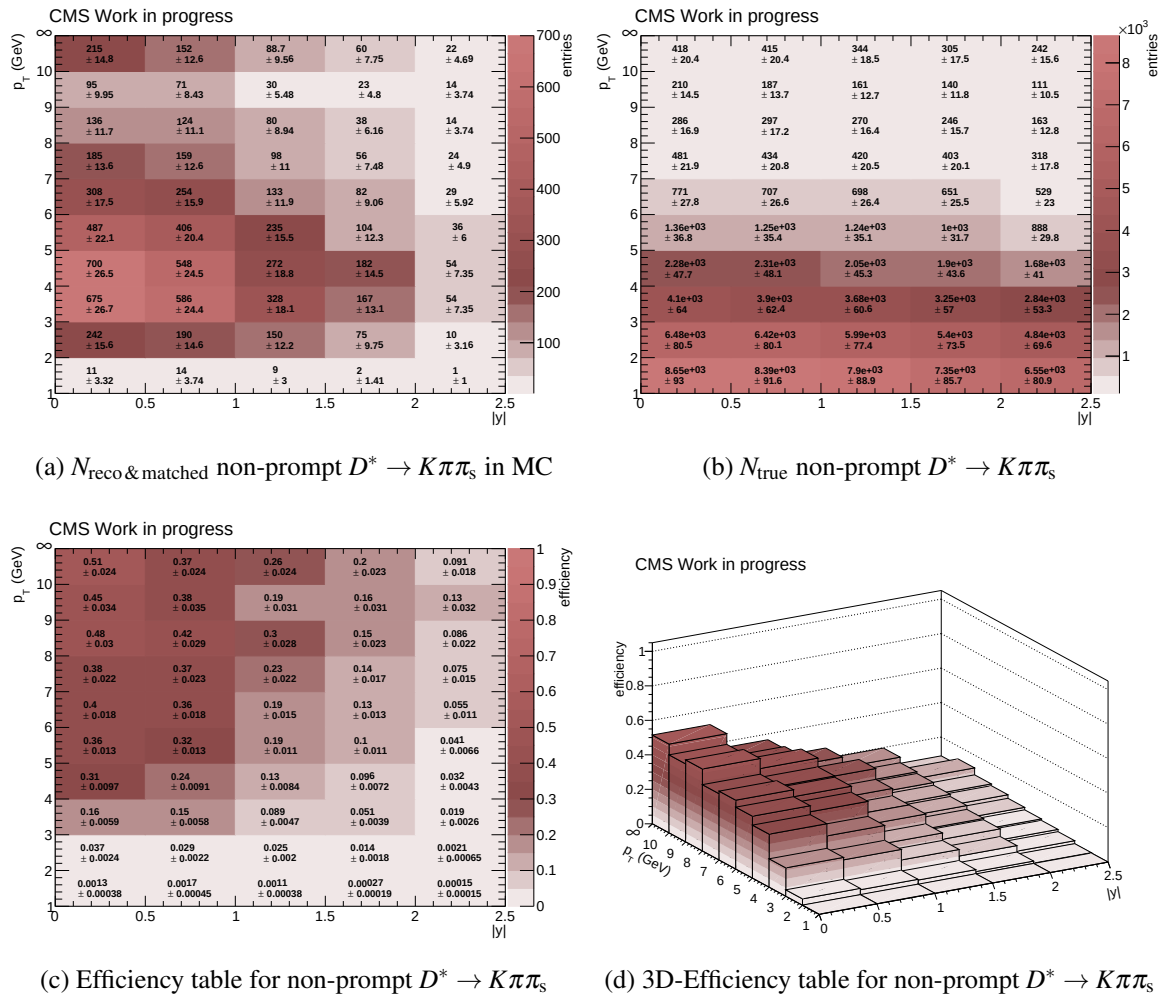


Fig. 6.10 Efficiency table and its statistical uncertainty for non-prompt $D^* \rightarrow K\pi\pi_s$ reconstruction in bins of p_T and y for $|y| < 2.5$, $p_T > 1.5$ GeV.

GeV and $|y| < 2.5$. Partial total cross sections are obtained in every bin of p_T and y by

$$\sigma_{pp \rightarrow D^* \rightarrow K\pi\pi_s} = \frac{N_{\text{signal}}}{\mathcal{L} \cdot \epsilon_{\text{efficiency}}} \quad (6.11)$$

which numerically correspond to cross sections differentially in p_T with a bin width of 1 GeV. Here, σ correspond to the cross section, N_{signal} to the number of the reconstructed $D^* \rightarrow K\pi\pi_s$ extracted via the subtraction method shown in Section 6.3, $\mathcal{L}$ to the integrated luminosity of the data Tab. 6.1 discussed in Section 6.1 and $\epsilon_{\text{efficiency}}$ to the efficiency discussed in Section 6.4.

The measurements of inclusive D^* cross sections are shown as function of p_T in the rapidity range $|y| < 0.5$ in Fig. 6.11. The distribution is shown differentially in p_T except for the last bin corresponding to the overflow bin. The measurements include only statistical uncertainties and are compared with theory predictions and existing overlapping results from ALICE [12]. For comparison, the shown results from ALICE include here only statistical uncertainties. The ratio in Fig. 6.11 shows the residuals between the theory prediction (or data) and the prediction by Pythia for inclusive D^* production. Since Pythia has LO + PS precision, these predictions are used only for shape comparison and scaled with a normalization factor of 0.8, as discussed in more detail in Section 6.8. The shape of Pythia for inclusive D^* production (c+b) agrees with the shape of this measurement. Theory predictions for prompt D^* production are obtained from FONLL as discussed in Chapter 5.

The measurements for the additional phase space are shown in the individual rapidity bins of $0.5 < |y| < 1$, $1 < |y| < 1.5$, $1.5 < |y| < 2$ and $2 < |y| < 2.5$ in Fig. 6.12. An overview of the entire measurement is shown in Fig. 6.13. Here, the measurements are compared with theory predictions and where it is possible with the existing overlapping results to LHCb [13] in the different rapidity ranges. The shown results from LHCb include only statistical uncertainties.

An overview of the entire measurement in the rapidity range from $0 < |y| < 2.5$ containing the comparisons with other results and theory predictions is shown in Fig. 6.13.

6.6 Separation of Charm and Beauty in D Meson Final States

Since the production of prompt and non-prompt D^* corresponds to charm and beauty production, respectively, the separation of charm and beauty in D meson final states is an indispensable ingredient for the determination of differential and total cross sections for

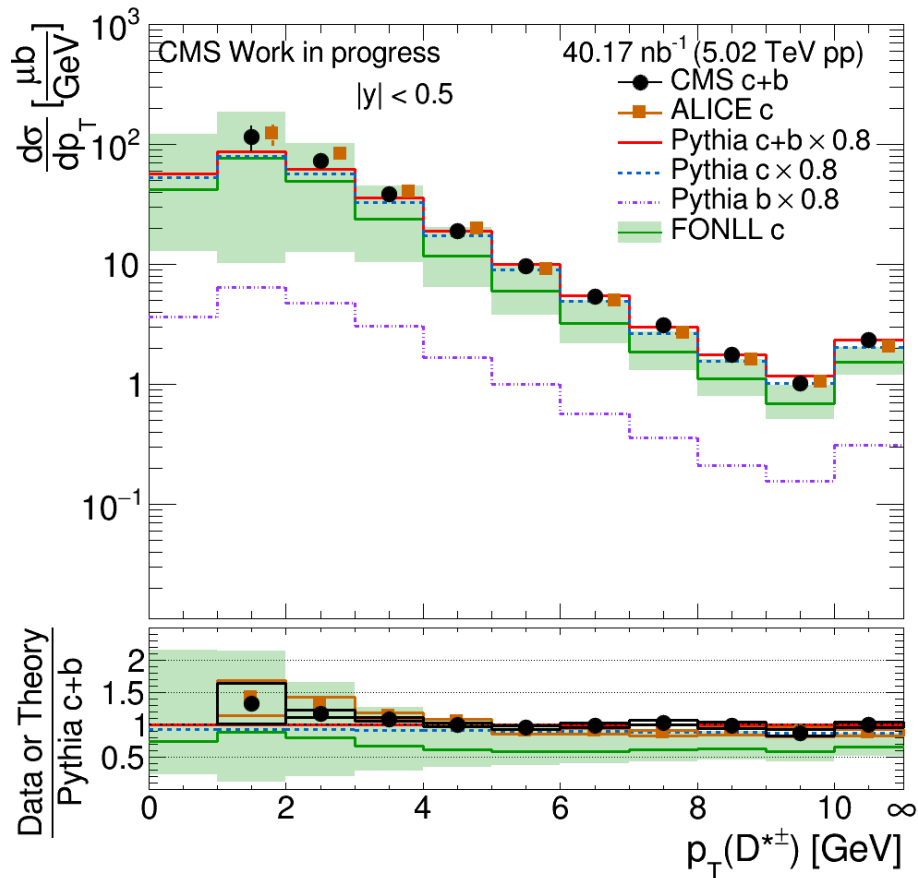


Fig. 6.11 Inclusive D^* cross sections as a function of p_T in the rapidity range $|y| < 0.5$. The measurements of this work for inclusive D^* production (c+b) are shown as black bullets, the dark orange rectangles correspond to the measurements from ALICE [12]. Predictions for cross sections by Pythia are shown in red for inclusive D^* and in blue (purple) for prompt (non-prompt) D^* . Predictions by Pythia are used only for shape comparisons and are scaled by 0.8. FONLL predictions for prompt D^* including an uncertainty band are shown in green.

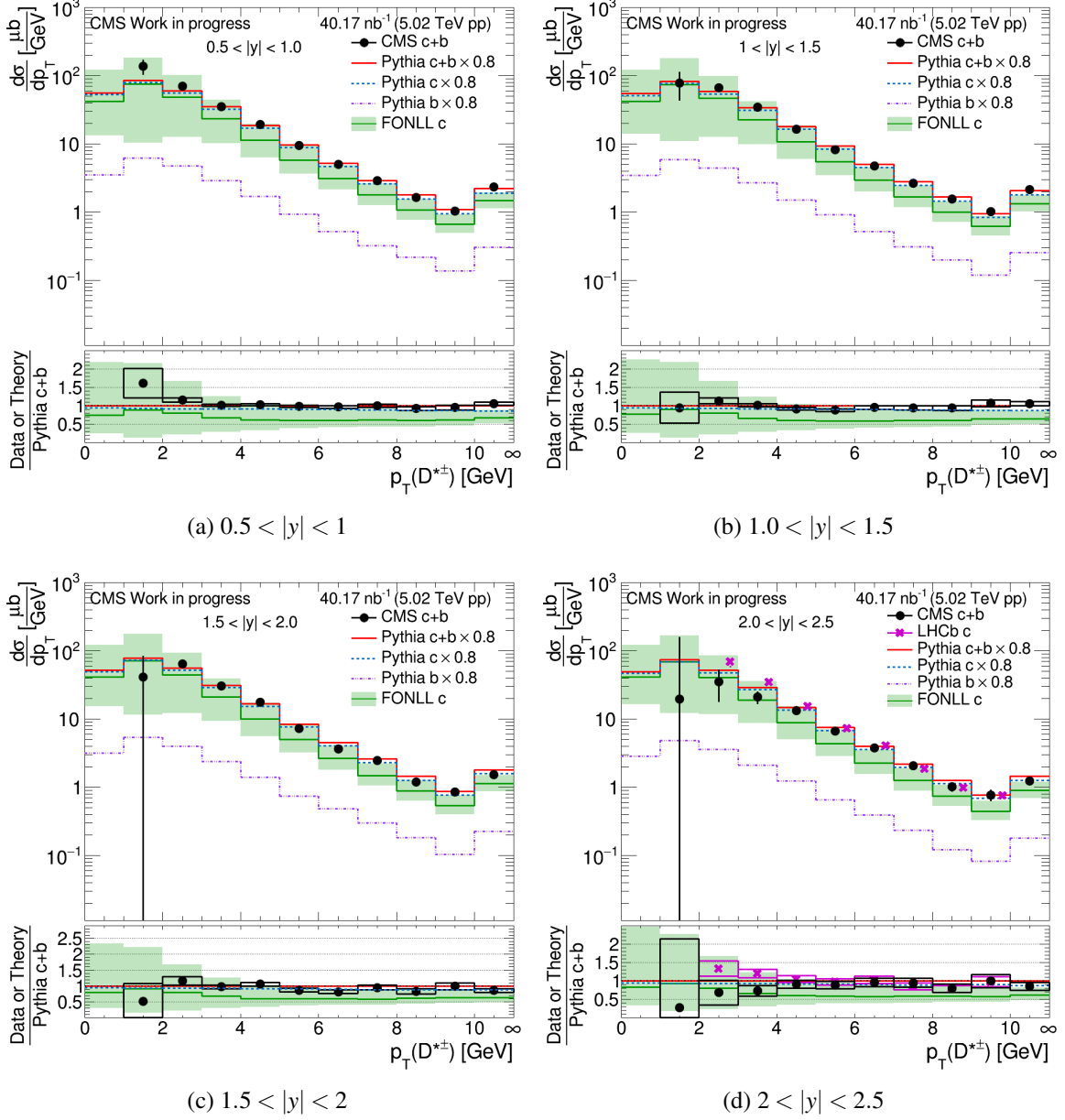


Fig. 6.12 Inclusive D^* cross sections as a function of p_T in different rapidity ranges ($|y| < 0.5$, $0.5 < |y| < 1$, $1 < |y| < 1.5$, $1.5 < |y| < 2$). The measurements of this work for inclusive D^* production ($c+b$) are shown as black bullets, the cyan rectangles correspond to the measurements from ALICE [12] and the purple cross bullets to the ones from LHCb [13]. Predictions by Pythia are shown in red for D^* in green (purple) for prompt (non-prompt) D^* . These are used only for shape comparisons and are scaled with 0.8. FONLL predictions for prompt D^* including an uncertainty band are shown in green.

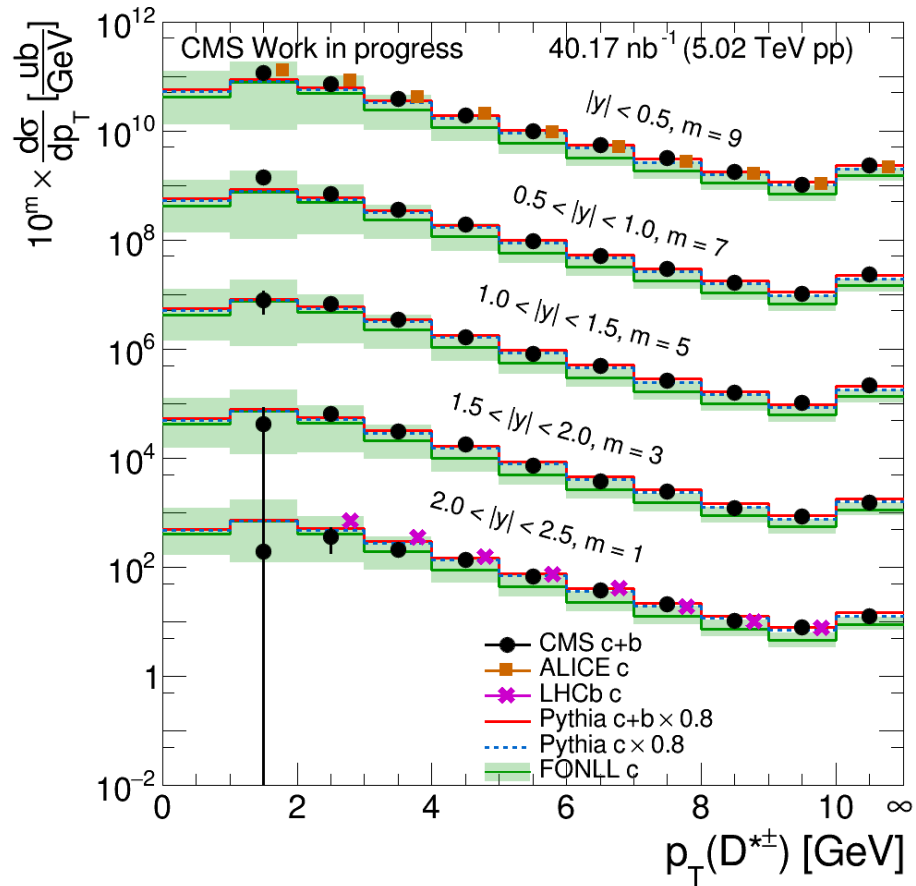


Fig. 6.13 Summary of the measurements of this work for inclusive D^* production. As black bullets the measurement of this work, as dark orange rectangles from ALICE [12] and as purple crosses to the one from LHCb [13]. Predictions by Pythia are shown in red for inclusive D^* and in green (purple) for prompt (non-prompt) D^* . FONLL predictions for prompt D^* including an uncertainty band are shown for prompt in green.

prompt and non-prompt D^* . A suitable observable for distinguishing between prompt and non-prompt has to be found, which in turn allows measuring directly the production of prompt or/and non-prompt D^* production. The method used here is based on the *distance of closest approach* (dca). It uses the different decay lengths of the D^0 and B hadrons listed in Tab. 5.2 and the different decay angle of prompt and non-prompt D^* . This procedure was also used in HIN [2, 3] to distinguish between prompt and non-prompt D^0 , where the dca of the D meson to the primary vertex is analyzed. A sketch illustrating the dca approach for a non-prompt D^* decay is shown in Fig. 6.14 and how the situation is changing in the case of prompt D^* production.

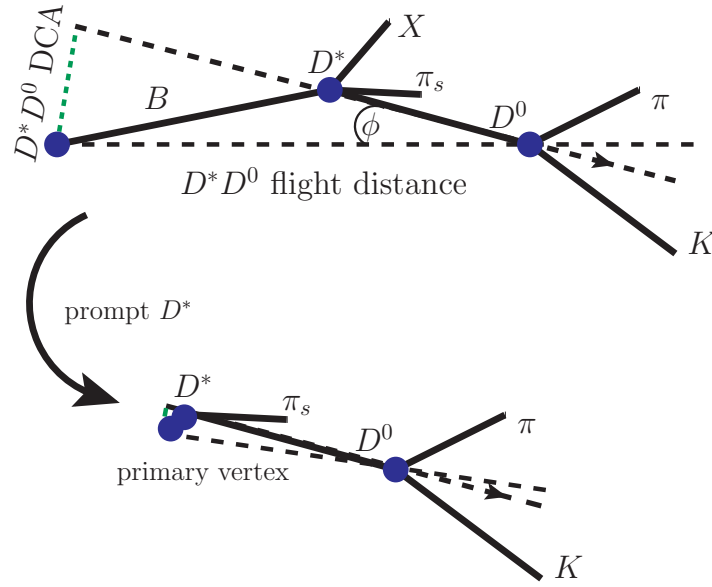


Fig. 6.14 A sketch illustrating the dca (distance of closest approach) for the production of a) non-prompt and b) prompt $D^* \rightarrow D^0 \pi_s \rightarrow K \pi \pi_s$. The dca correspond to the product of the flight distance, shown in the sketch, of the D^0 with respect to the PV and the angle ϕ between the momentum vector of the D^0 and the total momentum vector of the decay products (K and π). In the case of prompt D^* production, the dca is smaller. This is illustrated with the lower cartoons where the D^* production vertex now corresponds to the primary vertex.

Since the CMS detector has a limited resolution, only a statistical and not an event by event separation of prompt from non-prompt D^* is possible. To determine the dca of the D mesons to the primary vertex, the flight distance of a D^0 with respect to the primary vertex and the angle ϕ between the momentum vector of the D^0 and the total momentum vector the decay products (K and π) are used. The product of the flight distance of the D^0 and the angle

ϕ corresponds to the parameter dca which is calculated by

$$dca_{D^*D^0} = D^*D^0 \text{ flight distance} \cdot \sin \phi. \quad (6.12)$$

For the Δm signal region, dca distributions from the D^0 are measured for the right and wrong charge combinations discussed in Section 6.3. Comparing the distance of closest approach

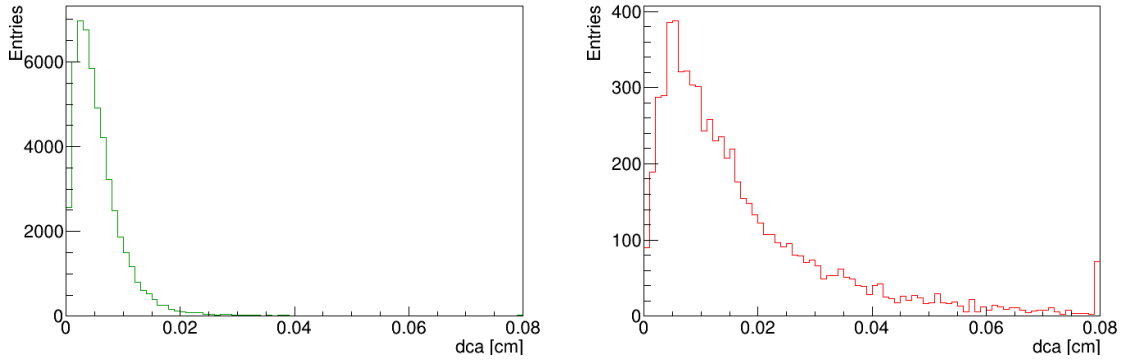


Fig. 6.15 On the left is an example for a reconstructed dca distribution for prompt D^* and on the right for non-prompt D^* . The distribution for D^* coming from b hadrons is harder compared to the prompt D^* .

distribution shown in Fig. 6.15, shows that the dca distribution for D^0 from D^* from b hadrons is harder compared to the prompt D^* . This difference in the shapes is the key to distinguish statistically between prompt and non-prompt D^* in D^* meson final states. In this way, the fractions of prompt and non-prompt D^* in D^* meson final states are determined as shown in Fig. 6.16 which lead to the measurements of charm and beauty cross sections presented in Section 6.8. For the fitting procedure an internal CMS tool called Higgs combine³ [14] is used which is based on RooFit⁴ functions. The tool uses multi-dimensional fits⁵ and likelihood based contours are performed too. This fit consists of two parameters, one for the fraction of prompt D^* and the other one for non-prompt D^* . The combinatorial background is scaled and constrained with the extracted SF discussed in Section 6.3 while the right charge combination corresponds to the distribution on which the fit is performed. Within this procedure a fit of each parameter is done separately while the other one is treated as unconstrained nuisance parameter [15]. The maximum and minimum of each parameter is determined in the 68% CL interval, according to a one-dimensional chi square. The statistical uncertainty of the MC samples of the charm and beauty templates is taken into account. The

³Combine uses different statistical techniques available inside RooFit/RooStats which are used widely within the CMS collaboration.

⁴RooFit is one of main classes of ROOT for building likelihood models.

⁵Via the module MultiDimFit

two fit parameters (c, b) for prompt D^* (charm, c) and non-prompt D^* (beauty, b) are anti correlated, since D^* final states consist only of charm and beauty.

After the fit, the charm and beauty contributions can be shown in Fig. 6.16 for the extracted signal in the higher p_T region. The higher p_T region will be used in the following as an example to illustrate how the charm beauty separation, its calculations and the extraction of the resulting cross sections are performed. The distribution for charm (green), beauty (red) and the combinatorial background scaled by the SF (gray) are fitted to the right charge combination and shown differentially in dca in Fig. 6.16. The combinatorial background corresponds to the scaled wrong charge combinations and the black bullets to the right charge combinations discussed in Section 6.3. The ratio in Fig. 6.16 shows the differences between the data (CMS black bullets) and the sum of the fitted distribution (charm, beauty, comb bg). The charm/beauty separation and dca fit are made in 13 dca bins consisting of two fit parameters (c, b) including their uncertainties (Δ_b, Δ_c). The determination of the signal for charm N_c (prompt D^*) and beauty N_b (non-prompt D^*) in the data can be extracted together with the obtained fit parameters c, b and their uncertainties Δ_b, Δ_c and their correlation Δ_{bc} .

For a concrete example, the determination of the individual signals is performed for the higher p_T region as a dedicated example in the following, to illustrate that the recombination of the signal for prompt and non-prompt D^* corresponds to the inclusive D^* signal extracted directly by the subtraction method. The correlation between charm and beauty is given by the covariance matrix

$$\text{cor}(c, b) = P_{\text{cor}} = \begin{pmatrix} \Delta_c^2 & \Delta_{bc} \\ \Delta_{bc} & \Delta_b^2 \end{pmatrix} = \begin{pmatrix} 6.27 \cdot 10^{-5} & -2.07 \cdot 10^{-4} \\ -2.07 \cdot 10^{-4} & 23.93 \cdot 10^{-4} \end{pmatrix}. \quad (6.13)$$

The extraction of the signal can be written as

$$\begin{aligned} N_{c+b}^{\text{fit}} &= N_c^{\text{fit}} + N_b^{\text{fit}} \\ &= (c \pm \Delta_c) \cdot N_c + (b \pm \Delta_b) \cdot N_b \\ &= (0.821 \pm 0.008) \cdot 52951 + (0.42 \pm 0.049) \cdot 6670 \\ &= (43472 \pm 423) + (2801 \pm 326) \\ &= 46274 \pm (\sqrt{(423^2 + 326^2 + 2c \cdot b \cdot N_b \cdot N_c \cdot \Delta_{bc})}) = 46274 \pm 298.2. \end{aligned} \quad (6.14)$$

Here, the uncertainties of the right charge combination and the wrong charge combination (combinatorial background) are taken into account and N_{c+b}^{fit} corresponds to the signal fitted to N_{c+b}^{sub} extracted by the data driven signal extraction method explained in Section 6.3. The combinatorial background is scaled here in this specific phase space region with SF= 1.097 and the extracted signal in data corresponds to $N_{c+b}^{\text{sub}} = 46202 \pm 307$, while the obtained

signal via the charm/beauty separation and dca fit is shown in Eq. 6.14. The recombination of the individual signals for charm and beauty is consistent within the uncertainties to the extracted signal (N_{c+b}^{sub}) for D^* without the charm beauty separation.

The determination of the fraction of prompt D^* (f_c , charm) and non-prompt D^* (f_b , beauty) is performed via

$$\begin{aligned} f_c &= \frac{C}{C+B}, \text{ with } C = c \cdot N_c \text{ and } \Delta_C = \Delta_c \cdot N_c, \\ \text{and } f_b &= \frac{B}{C+B}, \text{ with } B = b \cdot N_b \text{ and } \Delta_B = \Delta_b \cdot N_b. \end{aligned} \quad (6.15)$$

The correlation between the parameters for charm and beauty are taken into account such that

$$\begin{aligned} \Delta_{f_{b,c}}^2 &\approx \left| \frac{\partial f}{\partial B} \right|^2 \Delta_B^2 + \left| \frac{\partial f}{\partial C} \right|^2 \Delta_C^2 + 2 \frac{\partial f}{\partial B} \frac{\partial f}{\partial C} \Delta_{BC} \\ \Delta_{f_{b,c}} &\approx \sqrt{\frac{C^2}{(C+B)^4} \cdot \Delta_B^2 + \frac{B^2}{(C+B)^4} \cdot \Delta_C^2 - \frac{2BC}{(C+B)^4} \cdot \rho \Delta_B \Delta_C}. \end{aligned} \quad (6.16)$$

Due to the charm beauty separation, all individual parameters are known such that the uncertainties for the fraction of prompt D^* (Δ_{f_c} , charm) and non-prompt D^* (Δ_{f_b} , beauty) in D^* final states are derived by the propagation of uncertainty. This leads in total to

$$\begin{aligned} f_b \pm \Delta_{f_b} &= 0.059 \pm 0.0066 \\ f_c \pm \Delta_{f_c} &= 0.939 \pm 0.0066. \end{aligned} \quad (6.17)$$

Here, the two fit parameters for prompt D^* and non-prompt D^* , respectively, are anti correlated and thus, their uncertainty for the individual fraction for prompt D^* and non-prompt are the same $\Delta_{f_b} = \Delta_{f_c} = 6.65 \cdot 10^{-3}$. The uncertainty Δ_{f_b} corresponds to $\approx 11\% = \frac{\Delta_{f_b}}{f_b}$ and is tried to be reduced further within this study as shown in Section 6.7. The calculated fractions and their uncertainties are shown for the higher and lower p_T region in the legend of Fig. 6.16a and Fig. 6.16b, respectively. The shown charm beauty separation plots include the information about the individual fractions ($f_c \pm \Delta_{f_c}, f_b \pm \Delta_{f_b}$) in their legends too.

Due to statistical reasons and for the purpose of comparisons with existing results, the charm and beauty separation is performed in three rapidity regions, namely in rapidity $|y| < 0.5, |y| < 1$ and $1 < |y| < 2$. All bins shown in the signal extraction table in Fig. 6.5 were merged accordingly in these bins of rapidity. The efficiency table for prompt and non-prompt D^* was recalculated for this purpose. The signal extraction for all these merged

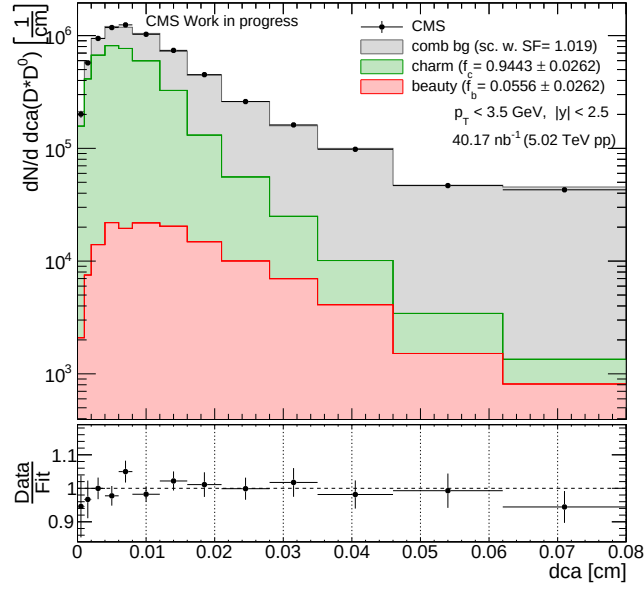
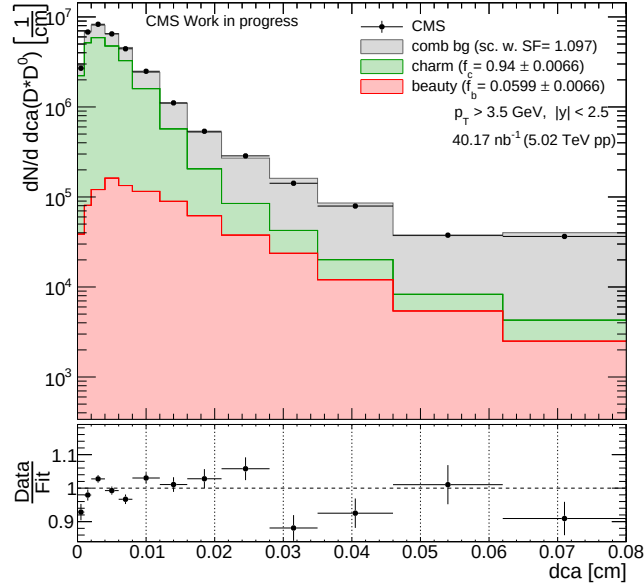
(a) Lower p_T region ($p_T < 3.5$ GeV) for $|y| < 2.5$ with dca.(b) Higher p_T region ($p_T > 3.5$ GeV) for $|y| < 2.5$ with dca.

Fig. 6.16 The separation of charm and beauty in D meson final states is shown in a) (in b)) in the lower (higher) p_T region for $|y| < 2.5$. The distribution for charm (green), beauty (red) and the combinatorial background (gray) are fitted to the right charge combination and shown differentially in dca. The combinatorial background corresponds to the wrong charge combination and is scaled with a scale factor SF coming from the subtraction method. The black bullets are labeled with CMS and correspond to the right charge combination. The ratio shows residuals between the data (CMS black bullets) and the sum of the fitted distribution (charm, beauty, combinatorial background). The measurement contains only statistical uncertainties.

bins can be found in Appendix C.1. The corresponding charm and beauty separation for these bins based on dca can be found in the Appendix C.2.

6.7 A Different Ansatz for the Charm Beauty Separation

This section shows the use of a different observable for the charm beauty separation. To reduce the uncertainty on the fraction of charm and especially for beauty, a set of different variables were studied within this work. For this purpose, correlation plots were studied showing different combinations of variables. The 2015 5 TeV $D^0 \rightarrow K\pi$ MC sample shown in Tab. 6.3 is used. In the following only the most suitable variable is shown while other correlation plots can be found in Appendix C.4.

A correlation between the decay length significance and the dca from the D^0 can be seen for charm (left) and beauty (right) in Fig. 6.17 where an elliptical (hyperbolic) correlation can be seen for beauty (charm). Due to this behavior, the product of the decay length significance with the dca from the D^0 coming from the D^* decay $D^* \rightarrow D^0 \pi_s$ is discussed in more detail in the further context. This product is given by

$$d_{\text{dca}} \cdot dl_{\text{sig}} = d_{\text{dca}} \cdot \frac{dl}{dl_{\text{err}}} . \quad (6.18)$$

The result of the charm and beauty separation performed with the product of the decay

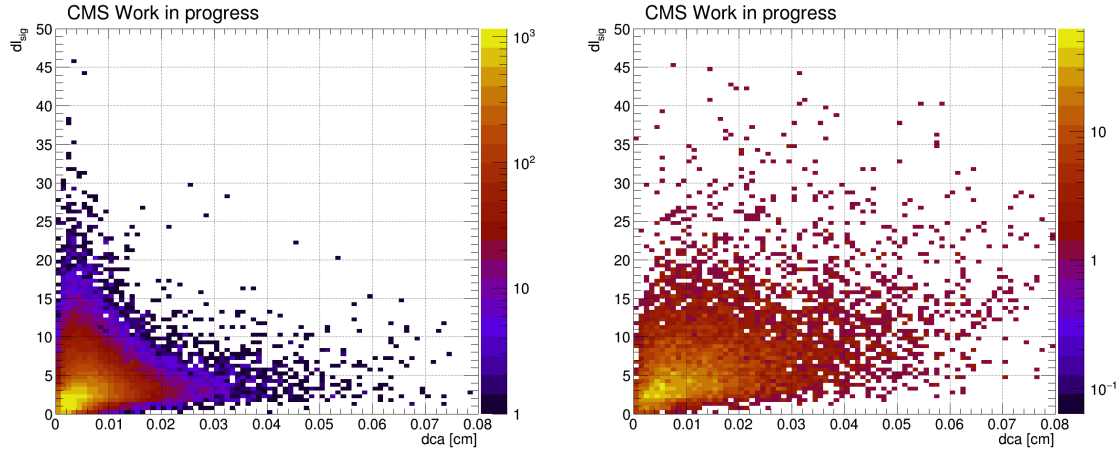


Fig. 6.17 Correlation plots of the decay length significance and the dca from the D^0 for charm (left) and beauty (right) for $p_T > 3.5$ GeV and $|y| < 2.5$. A hyperbolic (elliptical) correlation can be seen for charm (beauty) on the left (right).

length significance and the dca from the D^0 is shown for the higher and lower p_T region in Fig. 6.18. The determination of the signal for charm N_c and beauty N_b in data and of the

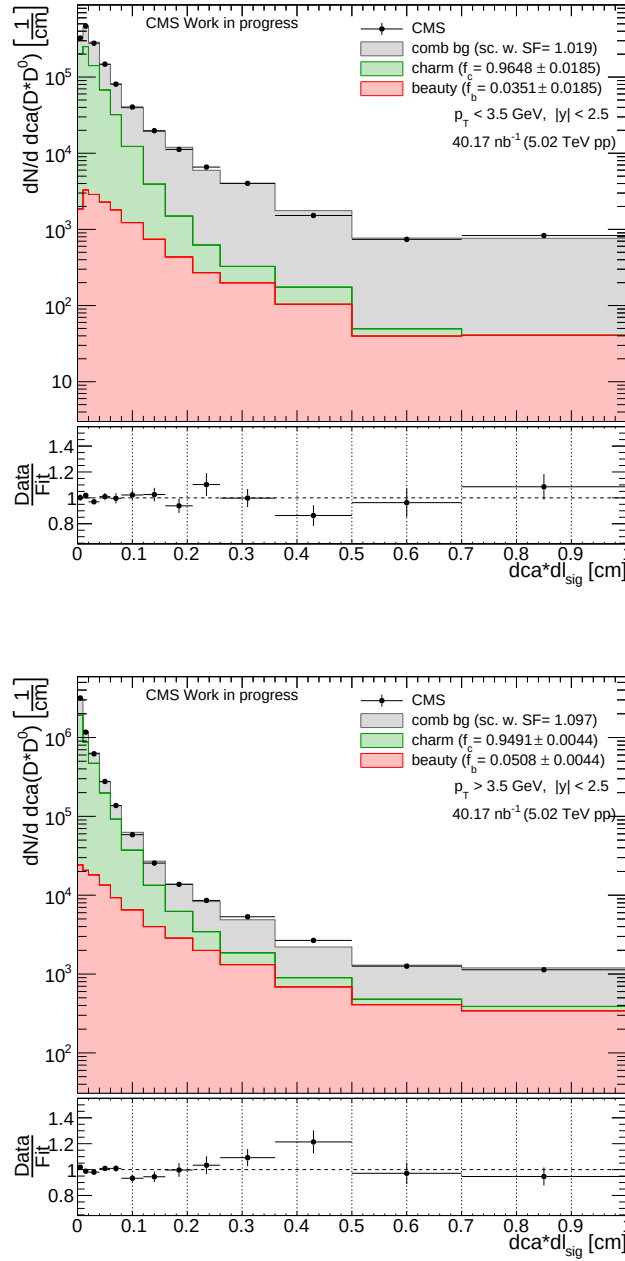


Fig. 6.18 The separation of charm and beauty in D^* meson final states performed via the method using the parameter $dca \cdot dl_{sig}$ is shown on the upper (lower) panel in the lower (higher) p_T region for $|y| < 2.5$. The distributions for charm (red), beauty (blue) and the combinatorial background (gray) are fitted to the right charge combination and shown differentially in the product of the decay length significance and the dca from the D^0 . The combinatorial background corresponds to the wrong charge combination and is scaled with a scale factor SF coming from the subtraction method. The black bullets are labeled with CMS and correspond to the right charge combination. The ratio shows the residuals between the data (CMS black bullets) and the sum of the fitted distribution (charm, beauty, combinatorial background).

fractions and their uncertainties for charm and beauty are calculated analogously according to Eqs. (6.14, 6.15, 6.16) with the obtained fit parameter c, b and their uncertainties $\Delta_{c,b}$. As an example, it follows the determination of the signal for charm N_c and beauty N_b in data for the higher p_T region. Following the procedure shown in Section 6.6 it can be written

$$\begin{aligned} N_{c+b} &= (c \pm \Delta_c) \cdot N_c + (b \pm \Delta_b) \cdot N_b \\ &= 46078.7 \pm 282.01 . \end{aligned} \quad (6.19)$$

The combinatorial background is scaled with $SF= 1.097$ and the extracted signal in data directly from the subtraction method corresponds to $N_{c+b} = 46202 \pm 307$. Comparing Eq. 6.19 with Eq. 6.14 shows that the uncertainty for the signal extraction is decreased by using the variable in Eq. 6.18 for the charm beauty separation. The calculation of the fractions and their uncertainties for charm and beauty leads to

$$\begin{aligned} f_b \pm \Delta_{f_b} &= 0.0508 \pm 0.0044 \\ f_c \pm \Delta_{f_b} &= 0.9492 \pm 0.0044 . \end{aligned} \quad (6.20)$$

Comparing the results from Eq. 6.20 with the one from Eq. 6.17 obtained via the method for the charm beauty separation using the dca , shows that the uncertainty Δ_{f_c} and as well Δ_{f_b} is decreased. Both methods are reasonably compatible. Using the parameter shown in Eq. 6.18 the uncertainty Δ_{f_b} correspond to $\approx 8.7\% = \frac{\Delta_{f_b}}{f_b}$ of the determined fraction for f_b . Based on $dca \cdot dl_{sig}$ is the charm and beauty separation performed for all bins and can be found in Appendix C.3.

6.8 Double Differential Prompt and Non-prompt D^* Cross Sections

This section shows the determination of the double differential prompt and non-prompt D^* cross sections. Comparisons with overlapping results from different experiments and with predictions by Pythia and FONLL are also made. For the comparison of the particular shapes for the predictions for inclusive, prompt and non-prompt D^* p_T distributions by Pythia, these LO+PS predictions will be scaled to the data. This procedure is shown in the following.

As references and for the purpose of shape comparison, unscaled predictions by Pythia obtained by using $\mathcal{L}_{eff}$ in Tab. 6.3 are shown in the following.

For prompt D^* it can be written that

$$\begin{aligned}\sigma_{pp \rightarrow D^*_{\text{non-prompt}}} &= \frac{N_{\text{true}}^{\text{non-prompt}}}{\mathcal{B}(D^* \rightarrow D^0 \pi) \cdot \mathcal{B}(D^0 \rightarrow K \pi) \cdot L_{\text{eff}}} \\ &= 30.77 \mu\text{b},\end{aligned}\quad (6.21)$$

while for non-prompt D^*

$$\begin{aligned}\sigma_{pp \rightarrow D^*_{\text{prompt}}} &= \frac{N_{\text{true}}^{\text{prompt}}}{\mathcal{B}(D^* \rightarrow D^0 \pi) \cdot \mathcal{B}(D^0 \rightarrow K \pi) \cdot L_{\text{eff}}} \\ &= 301.09 \mu\text{b}.\end{aligned}\quad (6.22)$$

Since the fractions of charm and beauty in D^* final states are measured in bins of p_T and $|y|$, the resulting differential cross sections for prompt and non-prompt D^* can be determined. The calculation of prompt and non-prompt D^* cross sections is performed for the dedicated example the higher p_T region and $|y| < 2.5$ and its separation via dca shown in Fig. 6.16a. The calculation for the efficiencies for charm and beauty in this phase region gives

$$\begin{aligned}\mathcal{E}_{\text{efficiency } b} &= \frac{N_{\text{reco \& matched}}^{\text{non-prompt}}}{N_{\text{true}}^{\text{non-prompt}}} = \frac{6670}{32783} = 0.203 \\ \mathcal{E}_{\text{efficiency } c} &= \frac{N_{\text{reco \& matched}}^{\text{prompt}}}{N_{\text{true}}^{\text{prompt}}} = \frac{52950}{320715} = 0.165.\end{aligned}\quad (6.23)$$

Here, $N_{\text{reco \& matched}}^{\text{non-prompt}}$ or $N_{\text{reco \& matched}}^{\text{prompt}}$ correspond to the signal shown previously in Fig. 6.6 and Fig. 6.7, respectively. Together with the $\mathcal{L}_{\text{int}}$ shown in Tab. 6.1 and the branching fractions $\mathcal{B}(D^* \rightarrow D^0 \pi)$ and $\mathcal{B}(D^0 \rightarrow K \pi)$ shown in Tab. 5.3, partial total cross sections for prompt and non-prompt D^* are obtained for the higher p_T region, $p_T > 3.5$ GeV and $|y| < 2.5$ by

$$\begin{aligned}\sigma_{pp \rightarrow D^*_{\text{non-prompt}} \rightarrow D^0 \pi \rightarrow K \pi \pi} &= \frac{N_{\text{signal}}^b}{L_{\text{int}} \cdot \mathcal{E}_b} = \frac{2801}{40.17 \cdot 0.203} = 342.7 \text{ nb} \\ \sigma_{pp \rightarrow D^*_{\text{non-prompt}}} &= 13009.9^{+1517.82}_{-1517.82} \text{ nb} = 13.01^{+1.51}_{-1.51} \mu\text{b} \\ \sigma_{pp \rightarrow D^*_{\text{prompt}} \rightarrow D^0 \pi \rightarrow K \pi \pi} &= \frac{N_{\text{signal}}^c}{L_{\text{int}} \cdot \mathcal{E}_c} = \frac{43472.7}{40.17 \cdot 0.165} = 6554.8 \text{ nb} \\ \sigma_{pp \rightarrow D^*_{\text{prompt}}} &= 248793^{+2424}_{-2424} \text{ nb} = 248.79^{+2.42}_{-2.42} \mu\text{b}.\end{aligned}\quad (6.24)$$

Here, N_{signal}^b or N_{signal}^c are measured by the charm/beauty separation (c.f. Fig. 6.16b) and their values correspond to the obtained one shown in Eq. 6.14.

The calculation of the scale parameters for the individual Pythia predictions is obtained by scaling the distribution from Pythia to the data. It is performed in the higher p_T region and the extracted scale factors are applied in the entire phase space to the individual distributions inclusive, prompt or non-prompt D^* . In this way, a shape comparison is obtained. The procedure follows for the inclusive D^* (c+b)

$$\frac{\sigma_{pp \rightarrow D^* \text{prompt}}^{\text{data}} + \sigma_{pp \rightarrow D^* \text{non-prompt}}^{\text{data}}}{\sigma_{pp \rightarrow D^* \text{prompt}}^{\text{Pythia}} + \sigma_{pp \rightarrow D^* \text{non-prompt}}^{\text{Pythia}}} = \frac{249 + 13}{301 + 31} = 0.79 \simeq 0.8, \quad (6.25)$$

and for prompt and non-prompt

$$\underbrace{\frac{\sigma_{pp \rightarrow D^* \text{prompt}}^{\text{data}}}{\sigma_{pp \rightarrow D^* \text{prompt}}^{\text{Pythia}}}}_{=S_c} = \frac{249}{301} = 0.83, \quad \underbrace{\frac{\sigma_{pp \rightarrow D^* \text{non-prompt}}^{\text{data}}}{\sigma_{pp \rightarrow D^* \text{non-prompt}}^{\text{Pythia}}}}_{=S_b} = \frac{13}{30.8} = 0.42. \quad (6.26)$$

These scale factors are used within this thesis and can be found in the individual results represented in differential cross sections plots for inclusive in Section 6.5 and for prompt and non-prompt D^* production in the following.

Prompt and non-prompt D^* production cross sections are measured via the charm/beauty separation using dca in each bin in p_T and rapidity y and the resulting differential cross sections are shown for the rapidity ranges $|y| < 0.5$, $|y| < 1$ and $1 < |y| < 2$ in Fig. 6.19. In these bins measurements of other experiments exist partially and due to the lack of statistic the bins in rapidity were partially merged. An overview of the whole measurement for $|y| < 2$ is shown in Fig. 6.19d.

The resulting cross sections measured via the charm beauty separation using the product of the dca and the decay length significance c.f. Eq. 6.18 are shown for the individual rapidity ranges in Fig. 6.20. Resulting cross sections measured via the charm beauty separation with the product of dca and the decay length significance of the D^0 are shown for $|y| < 0.5$, $|y| < 1$ and $1 < |y| < 2$ in Fig. 6.20. An overview of the entire measurement for $|y| < 2$ is given as Fig. 6.20d.

The measurement of prompt D^0 in [2] is rescaled to D^* predictions via the multiplication of the ratio of the fragmentation fractions of D^0 ($ff_{D^0} = 0.542$) and D^* ($ff_{D^*} = 0.236$) discussed in Chapter 5. The results of non-prompt [3] D^* production are rescaled by the ratio of the FONLL prediction for non-prompt D^* over D^0 , while this is performed

per bin in p_T . Predictions by Pythia are scaled with $S_c = 0.8$ and $S_c = 0.4$ according to the study presented above. Theory predictions for prompt (non-prompt) D^* including an uncertainty band are obtained by FONLL[16]. The two ratios show the residuals between the measurements (theory predictions) of prompt and non-prompt D^* production and the corresponding prediction by Pythia. For the purpose of comparisons, the shown results of the other experiments include only statistical uncertainties. An agreement with the results from ALICE and CMS in the comparable rapidity regions is obtained. The measurement of this work is in agreement with the prediction by FONLL (NLO + NLL approximation) and the shapes from Pythia.

The measured cross sections within this work are shown in Tabs. (6.7, 6.8, 6.9) for the method dca and in Tabs. (6.10, 6.11, 6.12) for $dca \cdot dl_{sig}$. Here no systematic uncertainties are shown and the statistical uncertainties do not include the uncertainty of the efficiency calculations. The statistical uncertainty contains the uncertainty derived by the subtraction method for $D^* \rightarrow K\pi\pi_s$. The results open opportunities for the combination with the measurements of LHCb [13] shown in Fig. 5.13 to extract the total charm production cross section with less theory extrapolation. Moreover, the results of this work can be combined with a measurement of prompt D^0 and other prompt D mesons to extract the charm fragmentation fractions. It can be even used as an input for a PDF analysis such as [17].

Table 6.7 Measured partial total cross sections in μb for prompt and non-prompt D^* production at 5 TeV in bins of p_T and in the rapidity range $|y| < 0.5$. The first bin is measured starting from 1.5 GeV and the last bin correspond to the overflow bin. All other bins have a bin width of 1 GeV. The charm beauty separation is performed with dca .

$ y < 0.5$		
$p_T[\text{GeV}]$	$\sigma_{D^*_{non-prompt}} \pm stat$	$\sigma_{D^*_{prompt}} \pm stat$
[1.5, 2]	0 ± 1.39	133.59 ± 29.22
[2, 3]	3.63 ± 2.08	72.24 ± 4.16
[3, 4]	1.16 ± 0.58	39.01 ± 1.23
[4, 5]	1.32 ± 0.25	18.09 ± 0.51
[5, 6]	0.57 ± 0.13	9.36 ± 0.29
[6, 7]	0.47 ± 0.12	5.02 ± 0.21
[7, 8]	0.27 ± 0.09	2.89 ± 0.14
[8, 9]	0.08 ± 0.03	1.72 ± 0.09
[9, 10]	0.11 ± 0.04	0.94 ± 0.07
[10, ∞]	0.17 ± 0.05	2.19 ± 0.09

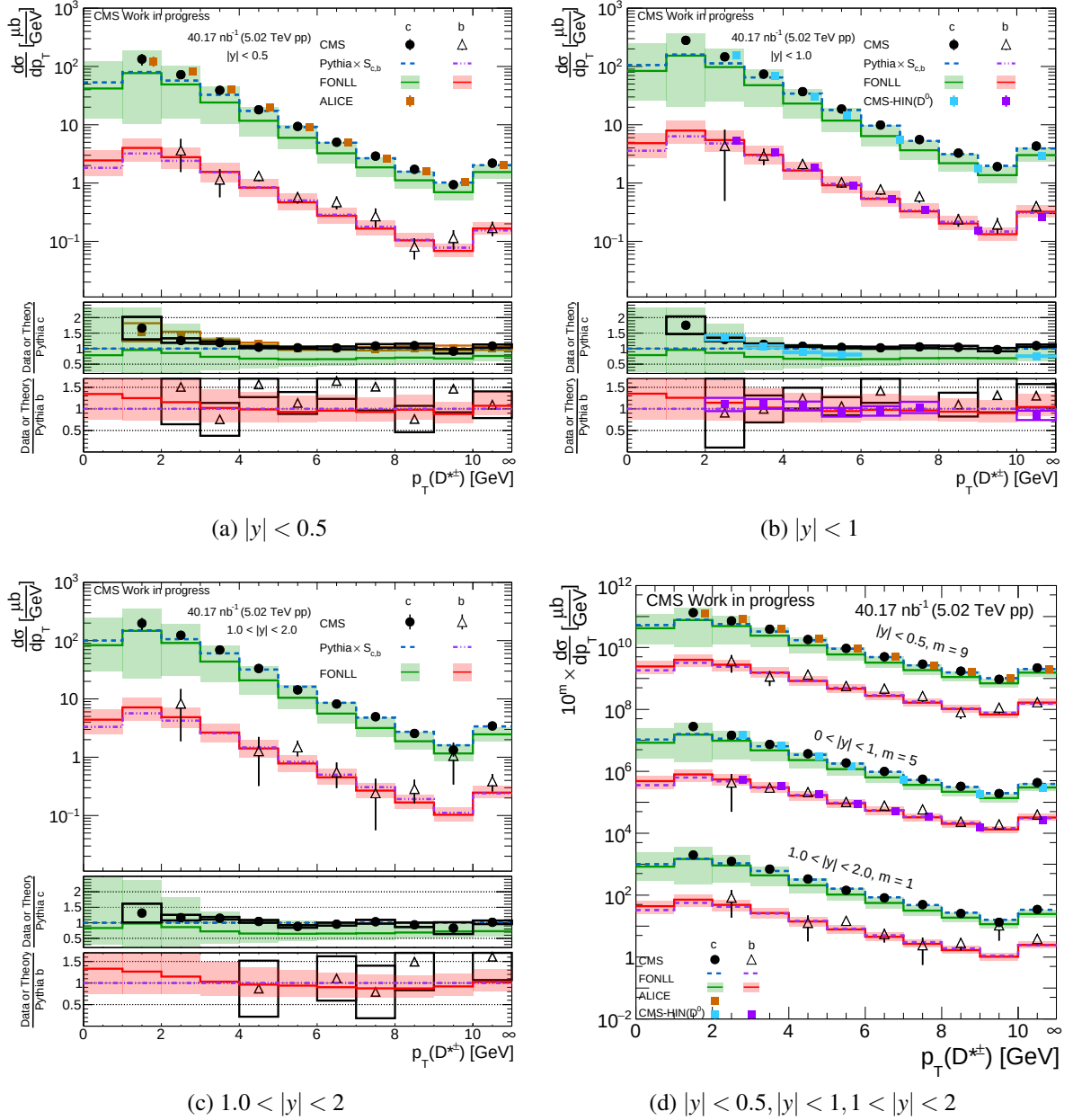


Fig. 6.19 Prompt and non-prompt D^* cross sections as a function of p_T in different rapidity ranges ($|y| < 0.5$, $|y| < 1$, $1 < |y| < 2$). These differential cross sections are determined using the parameter dca for the charm beauty separation. The measurement of this work for prompt (non-prompt) D^* is illustrated with black circle (triangle) bullets, the purple and light blue rectangle symbols to the one from HIN. The measurement from CMS-HIN of prompt [2] and non-prompt [3] D^0 is rescaled to D^* via the multiplication of the ratio of the fragmentation fractions of D^0 and D^* ($ff_{D^0} = 0.542$, $ff_{D^*} = 0.236$). Theory predictions for cross sections by Pythia are shown in blue(purple) for prompt (non-prompt) D^* . These are used only for shape comparison and scaled with $S_c = 0.8$ and $S_b = 0.4$, respectively. FONLL predictions including an uncertainty band for prompt (non-prompt) D^* are shown in green (red). The ratio shows the deviation between the measurement from the data of prompt and non-prompt D^* and the prediction by Pythia. In addition, it shows the deviation between the theory prediction by FONLL and the prediction by Pythia.

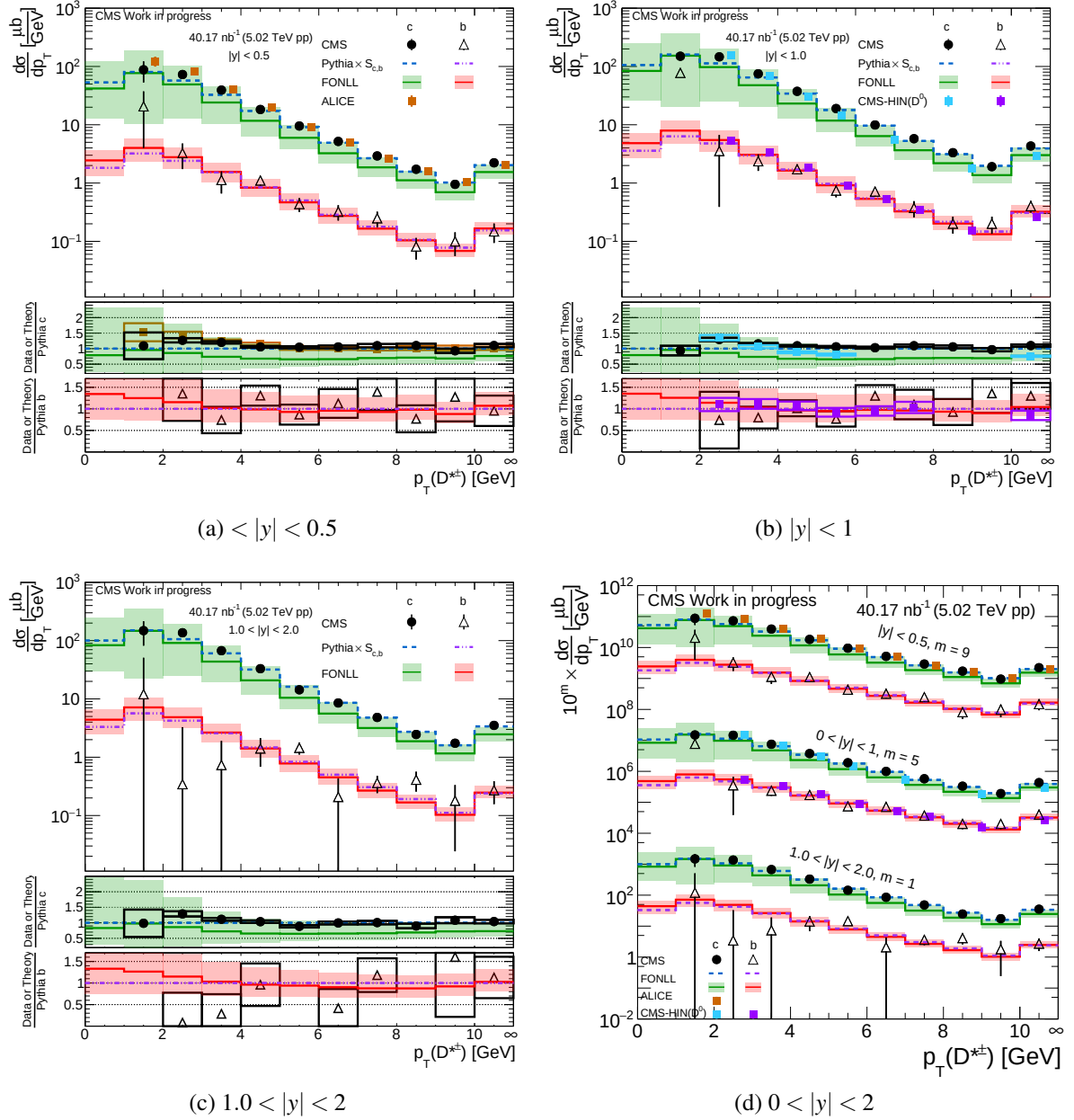


Fig. 6.20 Prompt and non-prompt D^* cross sections as a function of p_T in different rapidity ranges ($|y| < 0.5$, $|y| < 1$, $1 < |y| < 2$). These differential cross sections are determined using the parameter $dca \cdot d\text{LSig}$ for the charm beauty separation. The measurement of this work for prompt (non-prompt) D^* is illustrated with black circle (triangle) bullets, the purple and light blue rectangle symbols to the one from HIN. The measurement from CMS-HIN of prompt [2] and non-prompt [3] D^0 is rescaled to D^* via the multiplication of the ratio of the fragmentation fractions of D^0 and D^* ($f_{D^0} = 0.542$, $f_{D^*} = 0.236$). Theory predictions for cross sections by Pythia are shown in blue(purple) for prompt (non-prompt) D^* . These are used only for shape comparison and scaled with $S_c = 0.8$ and $S_b = 0.4$, respectively. FONLL predictions including an uncertainty band for prompt (non-prompt) D^* are shown in green (red). The ratio shows the deviation between the measurement from the data of prompt and non-prompt D^* and the prediction by Pythia. In addition, it shows the deviation between the theory prediction by FONLL and the prediction by Pythia.

Table 6.8 Measured partial total cross sections in μb for prompt and non-prompt D^* production at 5 TeV in bins of p_T and in the rapidity range $|y| < 1$. The first bin is measured starting from 1.5 GeV and the last bin correspond to the overflow bin. All bins have a bin width from 1 GeV. The charm beauty separation is performed with dca.

$ y < 1$		
$p_T[\text{GeV}]$	$\sigma_{D^*_{\text{non-prompt}}} \pm \text{stat}$	$\sigma_{D^*_{\text{prompt}}} \pm \text{stat}$
[1.5, 2]	0 ± 7.49	280.29 ± 45.39
[2, 3]	4.35 ± 3.86	146.25 ± 6.40
[3, 4]	2.96 ± 0.93	73.89 ± 1.82
[4, 5]	2.12 ± 0.41	36.96 ± 0.78
[5, 6]	1.03 ± 0.20	18.64 ± 0.44
[6, 7]	0.78 ± 0.15	9.89 ± 0.29
[7, 8]	0.59 ± 0.13	5.56 ± 0.19
[8, 9]	0.24 ± 0.06	3.27 ± 0.13
[9, 10]	0.19 ± 0.06	1.91 ± 0.10
[10, ∞]	0.40 ± 0.08	4.31 ± 0.14

Table 6.9 Measured partial total cross sections in μb for prompt and non-prompt D^* production at 5 TeV in bins of p_T and in the rapidity range $1 < |y| < 2$. The first bin is measured starting from 1.5 GeV and the last bin correspond to the overflow bin. All bins have a bin width from 1 GeV. The charm beauty separation is performed with dca.

$1 < y < 2$		
$p_T[\text{GeV}]$	$\sigma_{D^*_{\text{non-prompt}}} \pm \text{stat}$	$\sigma_{D^*_{\text{prompt}}} \pm \text{stat}$
[1.5, 2]	0 ± 12.18	197.96 ± 45.69
[2, 3]	8.33 ± 6.45	123.53 ± 9.45
[3, 4]	0 ± 0.89	69.29 ± 2.49
[4, 5]	1.27 ± 0.95	33.26 ± 1.36
[5, 6]	1.49 ± 0.45	14.28 ± 0.65
[6, 7]	0.55 ± 0.26	8.19 ± 0.39
[7, 8]	0.24 ± 0.19	4.95 ± 0.28
[8, 9]	0.26 ± 0.13	2.56 ± 0.19
[9, 10]	1.06 ± 0.72	1.33 ± 0.30
[10, ∞]	0.39 ± 0.13	3.44 ± 0.19

Table 6.10 Measured partial total cross sections in μb for prompt and non-prompt D^* production at 5 TeV in bins of p_T and in the rapidity range $|y| < 0.5$. The first bin is measured starting from 1.5 GeV and the last bin correspond to the overflow bin. All bins have a bin width from 1 GeV. The charm beauty separation is performed with $dca \cdot dl_{sig}$.

$ y < 0.5$		
$p_T [\text{GeV}]$	$\sigma_{D^*_{non-prompt}} \pm stat$	$\sigma_{D^*_{prompt}} \pm stat$
[1.5, 2]	20.67 ± 16.69	87.76 ± 34.60
[2, 3]	3.25 ± 1.51	72.69 ± 4.00971
[3, 4]	1.13 ± 0.47	39.30 ± 1.20733
[4, 5]	1.10 ± 0.19	18.26 ± 0.46
[5, 6]	0.44 ± 0.12	9.51 ± 0.30
[6, 7]	0.32 ± 0.09	5.17 ± 0.19
[7, 8]	0.25 ± 0.07	2.91 ± 0.13
[8, 9]	0.08 ± 0.03	1.72 ± 0.09
[9, 10]	0.10 ± 0.04	0.95 ± 0.07
[10, ∞]	0.15 ± 0.05	2.22 ± 0.10

Table 6.11 Measured partial total cross sections in μb for prompt and non-prompt D^* production at 5 TeV in bins of p_T and in the rapidity range $|y| < 1$. The first bin is measured starting from 1.5 GeV and the last bin correspond to the overflow bin. All bins have a bin width from 1 GeV. The charm beauty separation is performed with $dca \cdot dl_{sig}$.

$ y < 1$		
$p_T [\text{GeV}]$	$\sigma_{D^*_{non-prompt}} \pm stat$	$\sigma_{D^*_{prompt}} \pm stat$
[1.5, 2]	77.51 ± 0.28	149.81 ± 24.63
[2, 3]	3.54 ± 3.17	146.25 ± 6.19
[3, 4]	2.37 ± 0.75	74.89 ± 1.78
[4, 5]	1.76 ± 0.31	37.44 ± 0.79
[5, 6]	0.74 ± 0.17	18.96 ± 0.45
[6, 7]	0.71 ± 0.13	9.91 ± 0.29
[7, 8]	0.37 ± 0.12	5.76 ± 0.20
[8, 9]	0.20 ± 0.06	3.32 ± 0.13
[9, 10]	0.20 ± 0.07	1.91 ± 0.10
[10, ∞]	0.40 ± 0.09	4.32 ± 0.14

Table 6.12 Measured partial total cross sections in μb for prompt and non-prompt D^* production at 5 TeV in bins of p_T and in the rapidity range $1 < |y| < 2$. The first bin is measured starting from 1.5 GeV and the last bin correspond to the overflow bin. All bins have a bin width from 1 GeV. The charm beauty separation is performed with $dca \cdot dl_{sig}$.

$1 < y < 2$		
$p_T[\text{GeV}]$	$\sigma_{D^*_{non-prompt}} \pm stat$	$\sigma_{D^*_{prompt}} \pm stat$
[1.5, 2]	12.04 ± 38.80	148.42 ± 66.43
[2, 3]	0.34 ± 2.93	136.36 ± 8.78
[3, 4]	0.73 ± 1.17	67.15 ± 2.77
[4, 5]	1.41 ± 0.73	32.98 ± 1.29
[5, 6]	1.45 ± 0.34	14.34 ± 0.62
[6, 7]	0.21 ± 0.22	8.53 ± 0.39
[7, 8]	0.36 ± 0.12	4.82 ± 0.26
[8, 9]	0.41 ± 0.15	2.45 ± 0.19
[9, 10]	0.18 ± 0.15	1.74 ± 0.15
[10, ∞]	0.27 ± 0.11	3.51 ± 0.19

6.9 Systematic Uncertainties

This section gives a brief overview of some of the systematic uncertainties, which need to be estimated and taken into account in further studies. Based on MC samples, for instance, as mentioned and shown in Section 6.1 on a MB sample, systematic uncertainties can be estimated. Based on data, as shown in Section 3.3.2, the signal extraction for D^* reconstruction uncertainties regarding tracker alignments effects can be considered to be negligible. These effects have no impact on the kinematics from the D^* and the reconstruction of vertices. Furthermore, systematic uncertainties of the parameters which entered the cross section determination in this thesis like the branching fraction $\mathcal{B}(D^* \rightarrow D^0 \pi_s)$ or $\mathcal{B}(D^0 \rightarrow K\pi)$ need to be taken into account. As mentioned Section 6.1.1, in the integrated luminosity of 40.17 nb^{-1} can be used with a $\pm 2.3\%$ uncertainty [6]. Several sources of systematic uncertainty might need a time intensive study to be investigated and estimated which did not fit in this thesis.

References

- [1] CMS Collaboration, “Measurement of the $t\bar{t}$ production cross section, the top quark mass, and the strong coupling constant using dilepton events in pp collisions at $\sqrt{s} = 13$ TeV”, *Eur. Phys. J. C* **79** (2019), no. 5, 368, 10.1140/epjc/s10052-019-6863-8, arXiv:1812.10505.
- [2] CMS Collaboration, “Nuclear modification factor of D^0 mesons in PbPb collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, *Phys. Lett. B* **782** (2018) 474–496, 10.1016/j.physletb.2018.05.074, arXiv:1708.04962.
- [3] CMS Collaboration, “Studies of Beauty Suppression via Nonprompt D^0 Mesons in Pb-Pb Collisions at $Q^2 = 4$ GeV²”, *Phys. Rev. Lett.* **123** (2019), no. 2, 022001, 10.1103/PhysRevLett.123.022001, arXiv:1810.11102.
- [4] M. Cacciari, S. Frixione, and P. Nason, “The p(T) spectrum in heavy flavor photoproduction”, *JHEP* **03** (2001) 006, 10.1088/1126-6708/2001/03/006, arXiv:hep-ph/0102134.
- [5] “BRIL Work Suite”.
<https://cms-service-lumi.web.cern.ch/cms-service-lumi/brilwsdoc.html>.
- [6] CMS Collaboration, “CMS Luminosity Calibration for the pp Reference Run at $\sqrt{s} = 5.02$ TeV”,.
- [7] “McM Home page”, <https://cms-pdmv.cern.ch/mcm>.
- [8] “McM Home page”,
<https://cms-pdmv.cern.ch/mcm/public/restapi/requests/get/BPH-pp502Fall15-00001>.
- [9] R. Frühwirth, “Application of Kalman filtering to track and vertex fitting”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **262** (1987), no. 2, 444–450,
[https://doi.org/10.1016/0168-9002\(87\)90887-4](https://doi.org/10.1016/0168-9002(87)90887-4).
- [10] N. Z. Jomhari, “Measurement of Total and Differential Charm Cross Sections at 7 TeV with the CMS Detector”, 2022. Presented 2022-03-16.
<https://cds.cern.ch/record/2805338>.
- [11] ZEUS Collaboration, “Measurement of charm fragmentation fractions in photoproduction at HERA”, *JHEP* **09** (2013) 058, 10.1007/JHEP09(2013)058, arXiv:1306.4862.
- [12] ALICE Collaboration, “Measurement of D^0 , D^+ , D^{*+} and D_s^+ production in pp collisions at $\sqrt{s} = 5.02$ TeV with ALICE”, *Eur. Phys. J. C* **79** (2019), no. 5, 388, 10.1140/epjc/s10052-019-6873-6, arXiv:1901.07979.
- [13] LHCb Collaboration, “Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 5$ TeV”, *JHEP* **06** (2017) 147, 10.1007/JHEP06(2017)147, arXiv:1610.02230.

-
- [14] “TWiki page: LHC Higgs Combination Group”,
<https://twiki.cern.ch/twiki/bin/view/LHCPhysics/HiggsCombination>.
 - [15] “CombineHarvester”, <https://cms-analysis.github.io/CombineHarvester/index.html>.
 - [16] M. Cacciari, M. Greco, and P. Nason, “The P(T) spectrum in heavy flavor hadroproduction”, *JHEP* **05** (1998) 007, 10.1088/1126-6708/1998/05/007, arXiv:hep-ph/9803400.
 - [17] PROSA Collaboration, “Improved constraints on parton distributions using LHCb, ALICE and HERA heavy-flavour measurements and implications for the predictions for prompt atmospheric-neutrino fluxes”, *JHEP* **04** (2020) 118, 10.1007/JHEP04(2020)118, arXiv:1911.13164.

Chapter 7

Conclusion and Outlook

The result of this thesis is part of a project to measure the total charm and beauty cross sections at different pp center of mass energies (0.9, 2.76, 5, 7, 8, and 13 TeV), such that the dependence of the cross section for the charm/beauty pair production on the center of mass energy is obtained with minimal theory extrapolation. The idea is to measure cross sections in the full phase space of prompt D mesons and D mesons from b hadron decays in bins in p_T and $|y|$ and to integrate those to get the total charm/beauty cross section. The focus of this work was to measure the D^* cross section double differentially at $\sqrt{s} = 5$ TeV.

The challenge to separate prompt from non-prompt D^* at the production threshold was approached by exploiting observables sensitive to the long lifetime of b hadrons. The D^* cross section is measured double differentially in bins of D^* rapidity ($|y| < 2.0$) and transverse momentum ($p_T > 1.5$ GeV). The results of this measurement are compared to and found to be consistent with, NLO + NLL prediction (FONLL) and other measurements. This work is the first measurement of D^* from charm and beauty in CMS at 5 TeV and covers the full accessible phase space. The phase space region $1 < |y| < 2$ is measured for the first time at the LHC.

A combination with LHCb ($|y| > 2.0$) results in almost full coverage of the phase space, thus allowing e.g. the extraction of the total charm cross section. A possible extension for a potential future analysis can be to measure the Yukawa coupling between the Higgs boson and the beauty quark.

The differential cross sections measured in this work can also be used as input which, together with e.g. the data from ALICE and HERA, constrain the gluon and charm quark components in PDF fits.

Appendix A

Track Based Alignment

This Appendix summarizes all information about the track based alignment approach which could not fit into the main chapter of this thesis.

A.1 Weak Modes

Global distortions (weak modes) to which Eq. 3.7 is not sensitive are studied. An illustration of weak modes are given with Fig. A.1.

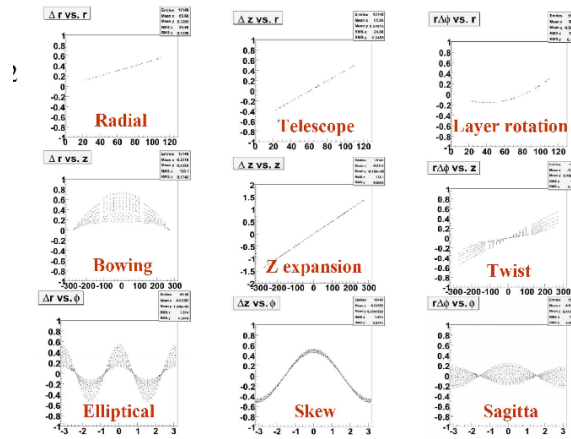


Fig. A.1 An illustration of all weak modes. This figure is taken from Reference [1].

A.2 Testing the Impact on Reconstructed Secondary Vertices due to the Improvement of the Alignment

Within this PhD, two different alignment eras for 2018 (prompt reconstruction and ultra legacy) are studied to test whether there is an impact on reconstructed secondary vertices due to the improvement of the alignment. Distributions for D mesons are shown in Fig. A.2 and Fig. A.3. Distributions for J/Ψ are shown in Fig. A.4.

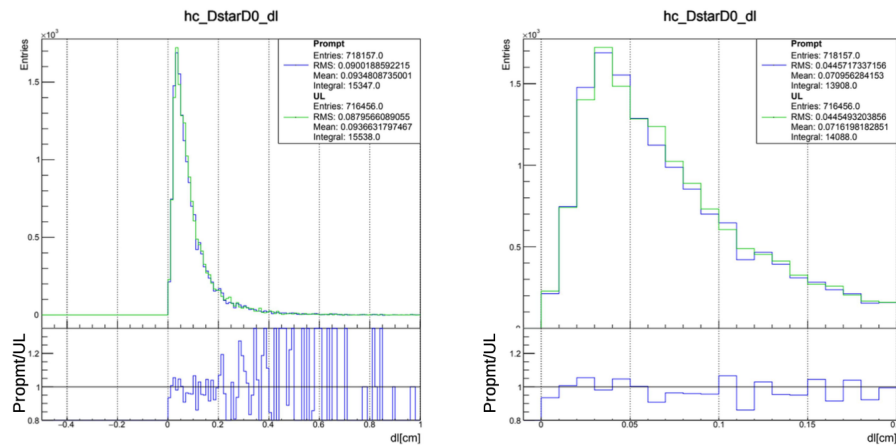


Fig. A.2 Decay length D meson distribution.

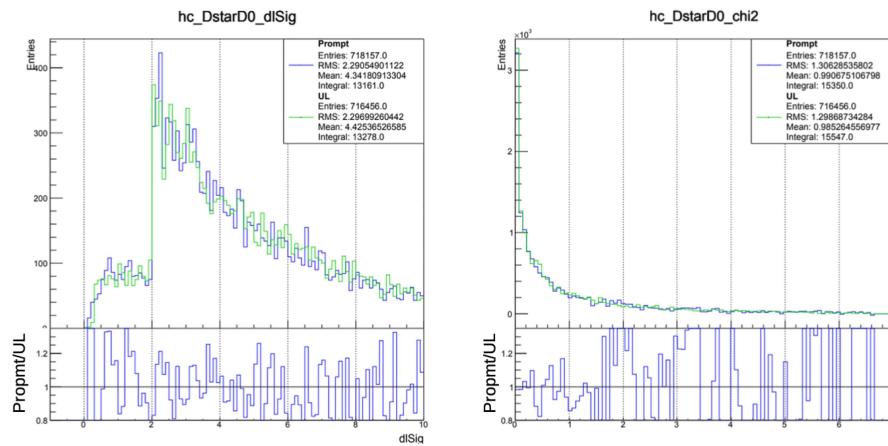


Fig. A.3 Decay length significance D meson distribution.

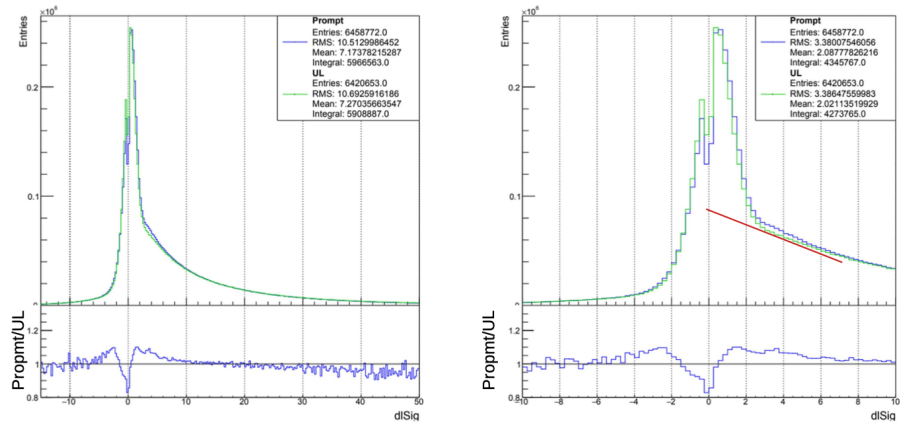


Fig. A.4 J/Ψ decay length significance distribution. Subtracting non prompt J/Ψ 's (shown here with the red line) lead to a more symmetric distribution for the UL alignment.

References

- [1] J. Draeger, “P o S (I C H E P 2 0 1 0) 5 0 4 The alignment of the CMS Silicon Tracker”, 2010.

Appendix B

Measurement of inclusive D^* Cross Sections in Bins of p_T and Rapidity

B.1 Effective Integrated Luminosity for pp Minimum Bias Data at 5 TeV

In the following the executed command is shown with its entire output. The used normtag list the selection of events where a luminosity measurement was taken. The delivered and recorded luminosity and a list of missing luminosity sections (LS) can be seen. The delivered luminosity is the one which is delivered by the LHC. Since sometimes the data acquisition or some detector subsystems are busy, the recorded luminosity by CMS does not correspond to the delivered one from the LHC. In bracket the LS of the used JSON and normtag can be seen which belongs to the run number in the beginning of each line of the output of the command. It can be seen that for some LS the normtag is not a superset from the used JSON which in turn means that for some data no luminosity measurement exists. The missing LS has been checked to contain LS with 100% dead-time. Thus, the final luminosity determination of 40.17 nb^{-1} is not affected by these LS [1].

```
brilcalc lumi --normtag /cvmfs/cms-bril.cern.ch/cms-lumi-pog/Normtags/normtag_PHYSICS.json
-i Cert_262081-262328_5TeV_PromptReco_Collisions15_25ns_JSON.txt
--hltpath HLT_LiMinimumBiasHF10R_part*_v1 -u /nb
This leads to:
#Sum delivered : 41.096511635
#Sum recorded : 40.172530301

Warning: problems found in merging -i and --normtag selections:
in run 262272 [[1, 232]] is not a superset of [[1, 241]]
in run 262273 [[1, 78]] is not a superset of [[1, 85]]
in run 262274 [[1, 1006]] is not a superset of [[1, 403], [406, 449], [452, 495], [499, 573],
[576, 596], [599, 1014]]
```

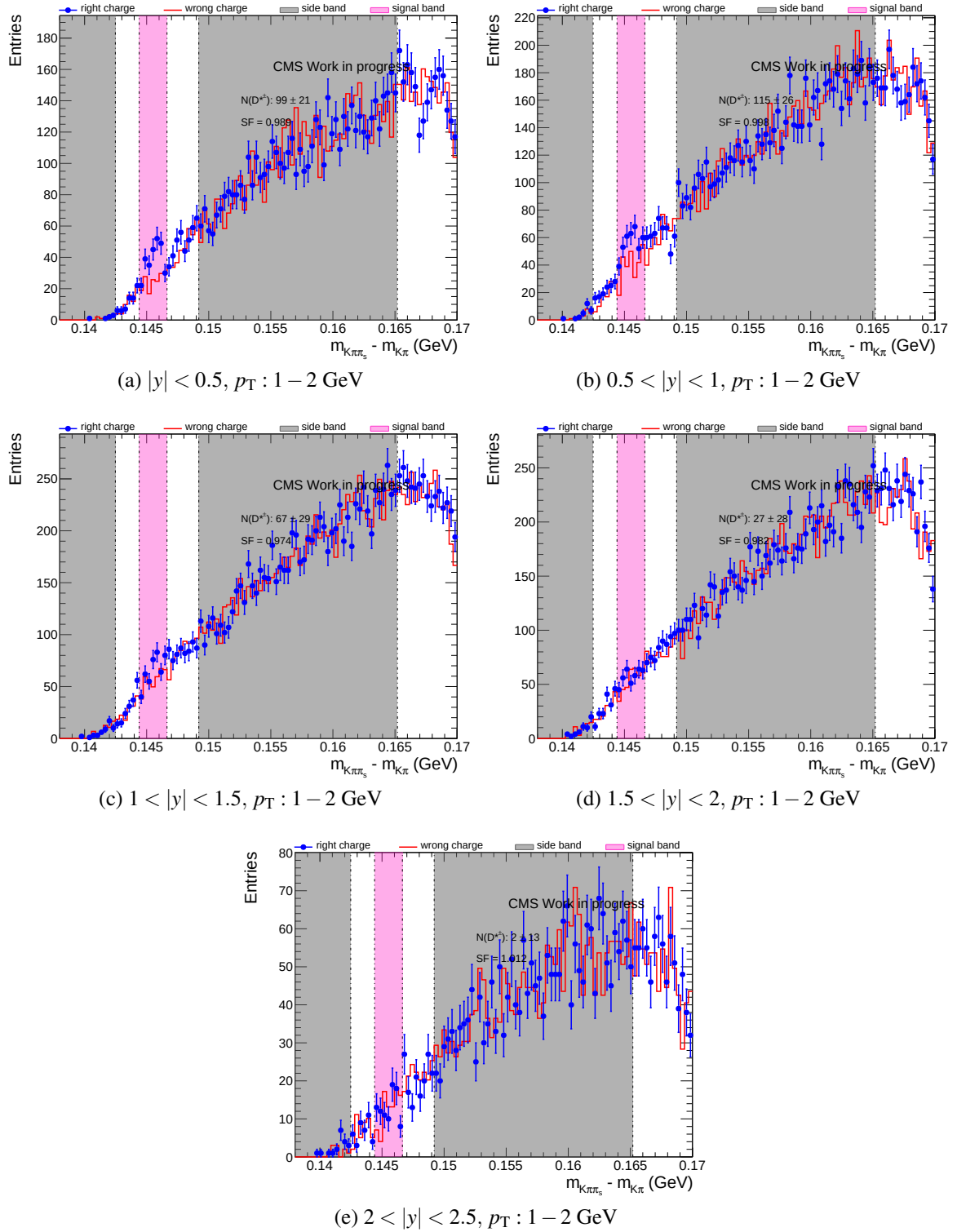
```

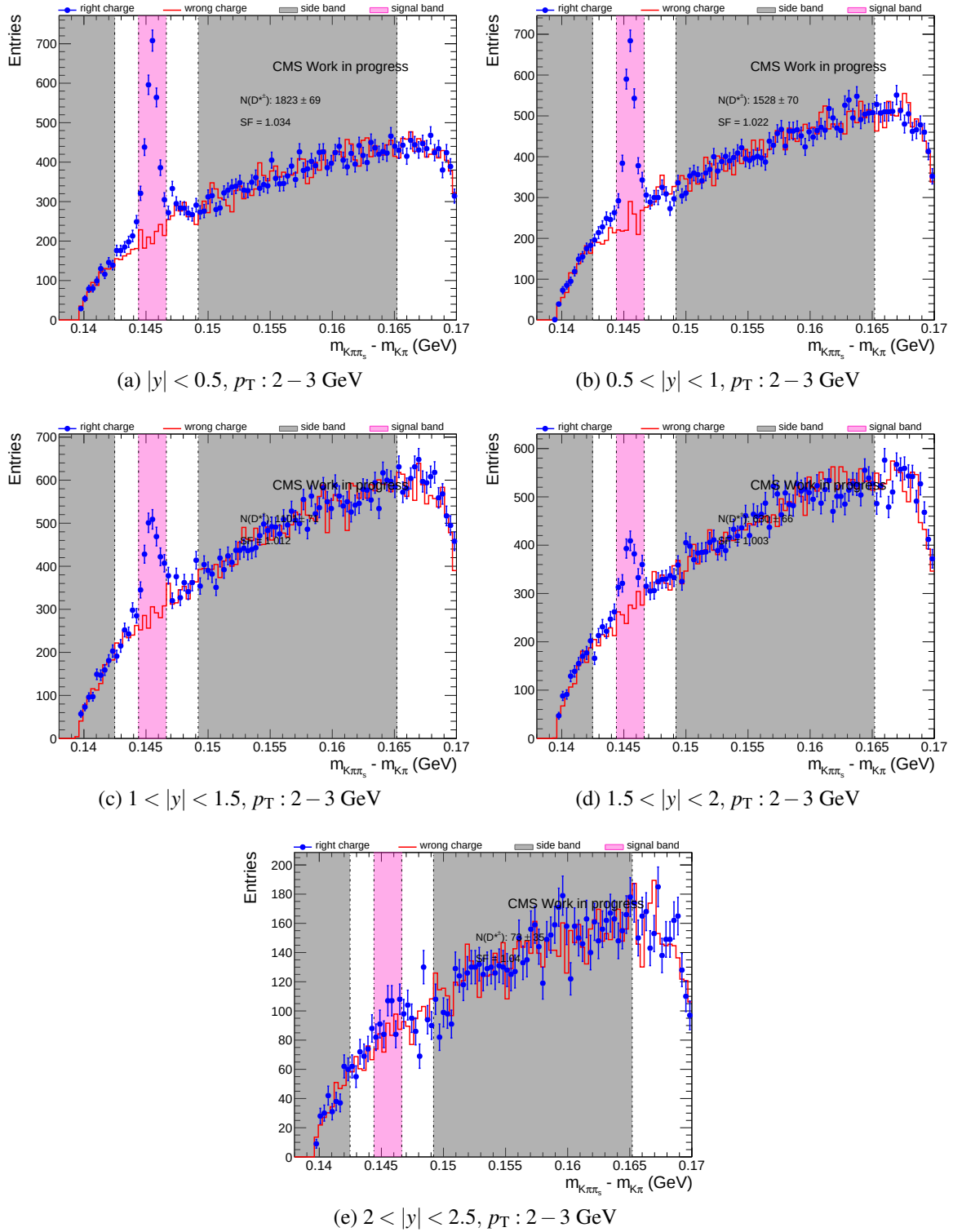
in run 262275 [[1, 78]] is not a superset of [[1, 87]]
in run 262248 [[27, 29]] is not a superset of [[26, 32], [35, 35]]
in run 262250 [[1, 92]] is not a superset of [[1, 32], [34, 93], [96, 96], [99, 100]]
in run 262327 [[1, 233]] is not a superset of [[1, 240]]
in run 262252 [[1, 233]] is not a superset of [[1, 241]]
in run 262253 [[1, 155]] is not a superset of [[1, 70], [73, 163]]
in run 262163 [[1, 235]] is not a superset of [[1, 117], [119, 242]]
in run 262165 [[1, 133]] is not a superset of [[1, 40], [45, 139]]
in run 262326 [[1, 78]] is not a superset of [[1, 63], [66, 69], [75, 86]]
in run 262167 [[1, 278]] is not a superset of [[1, 170], [173, 280]]
in run 262235 [[24, 24], [26, 379]] is not a superset of [[24, 287], [289, 379]]
in run 262173 [[1, 182]] is not a superset of [[1, 190]]
in run 262270 [[28, 79]] is not a superset of [[27, 87]]
in run 262271 [[1, 342]] is not a superset of [[1, 306], [309, 329], [335, 344]]

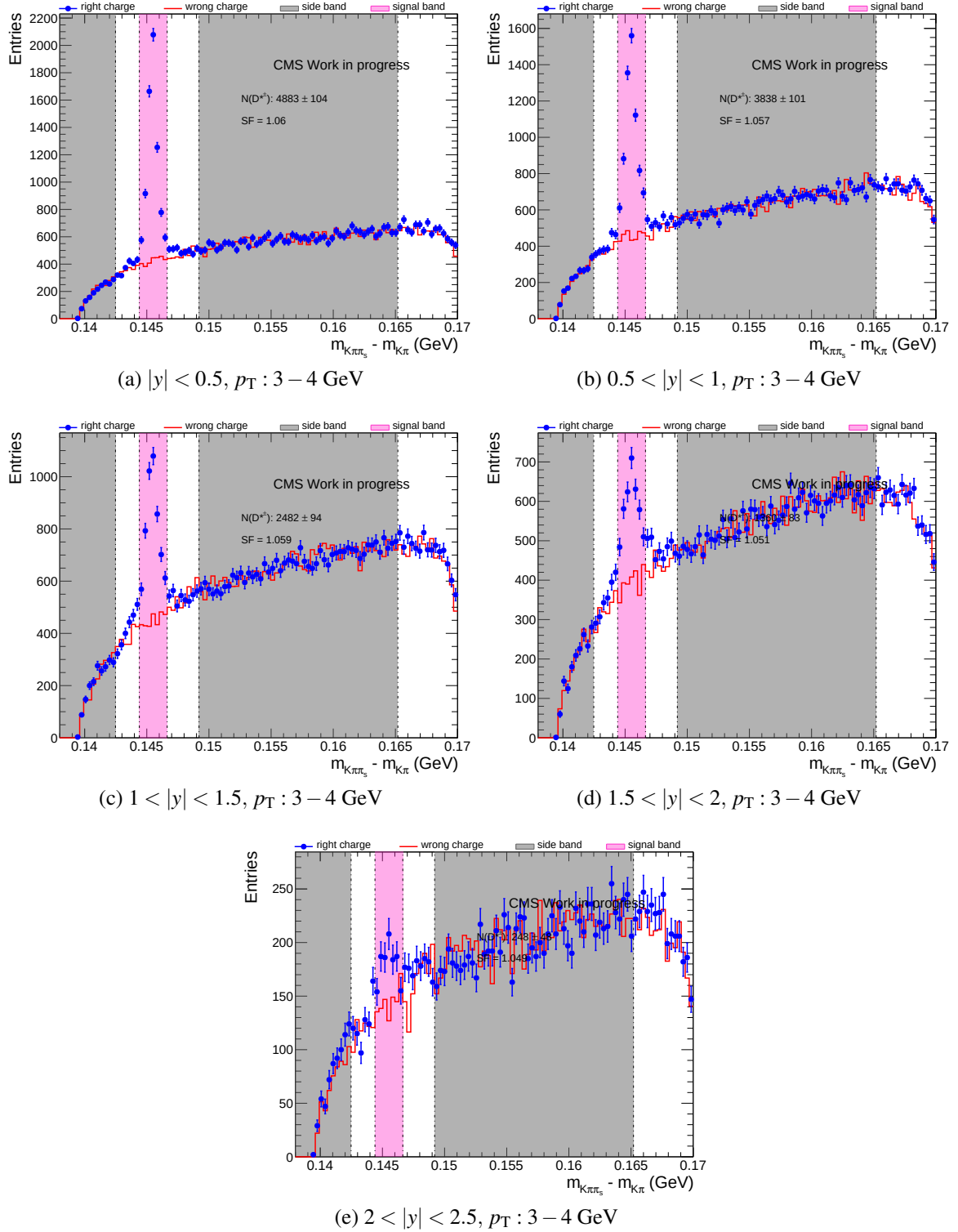
```

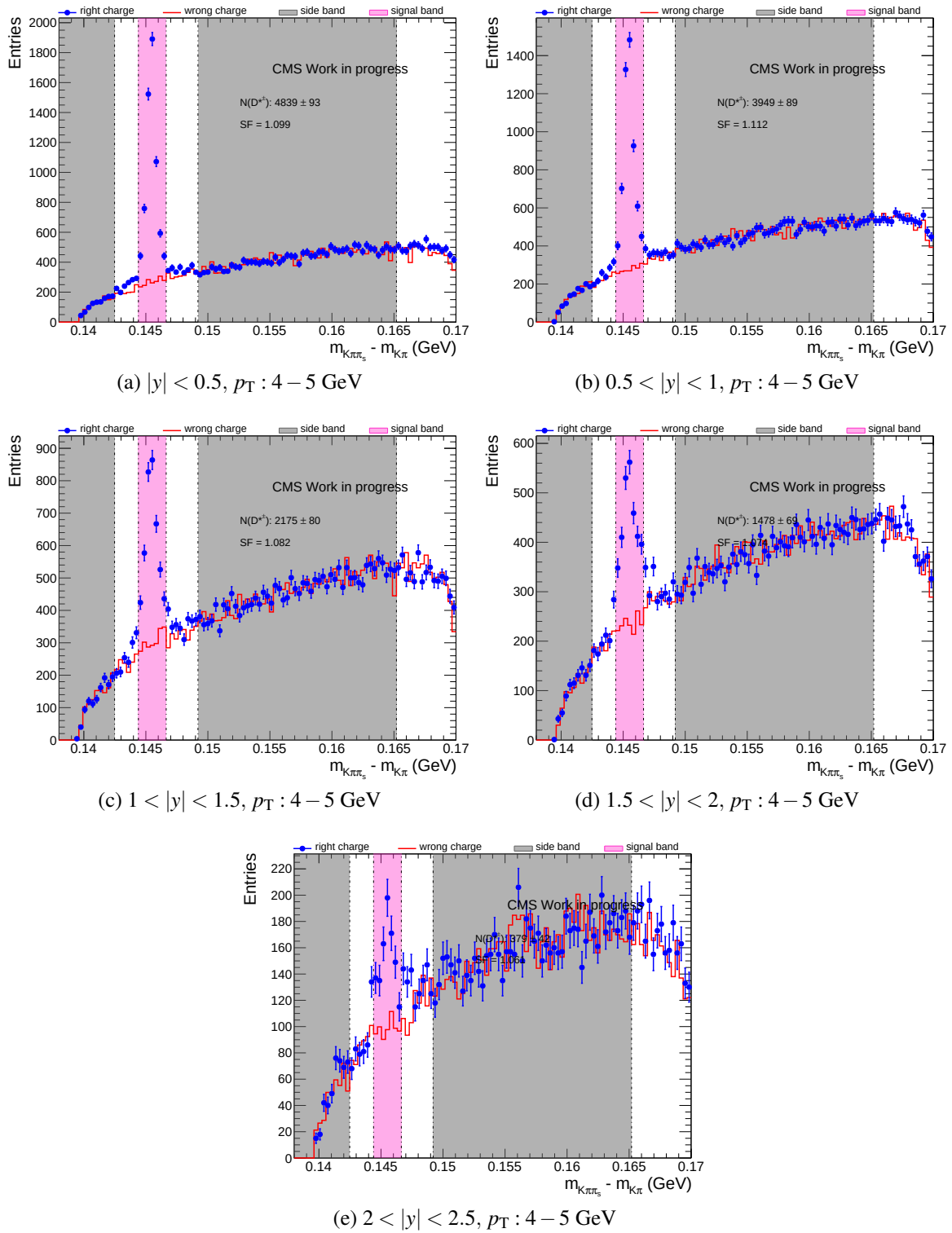
B.2 Signal Extraction for D^* in Bins of p_T and Rapidity

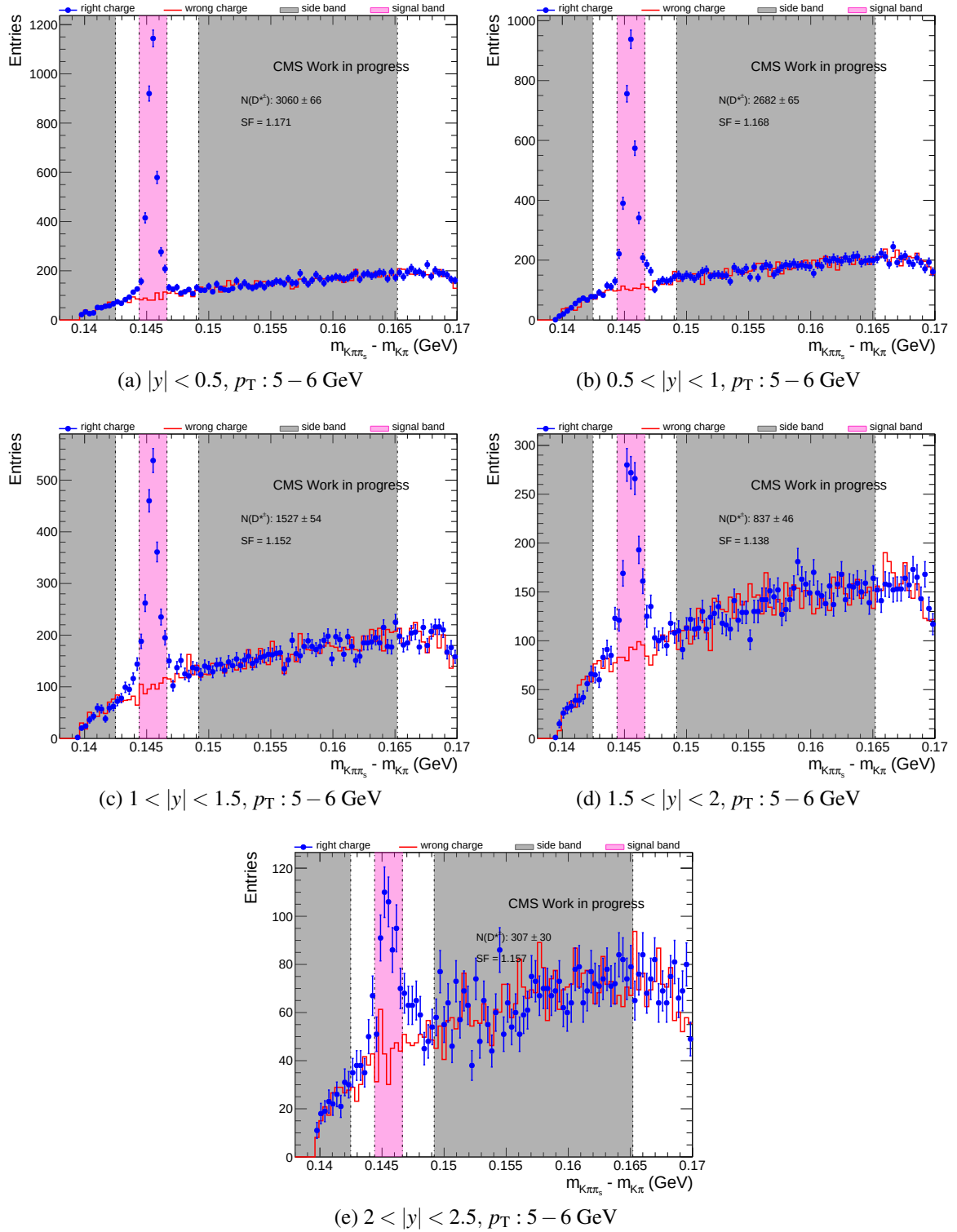
In the following signal extraction plots in the bins of p_T and rapidity $|y|$ are shown for D^* .

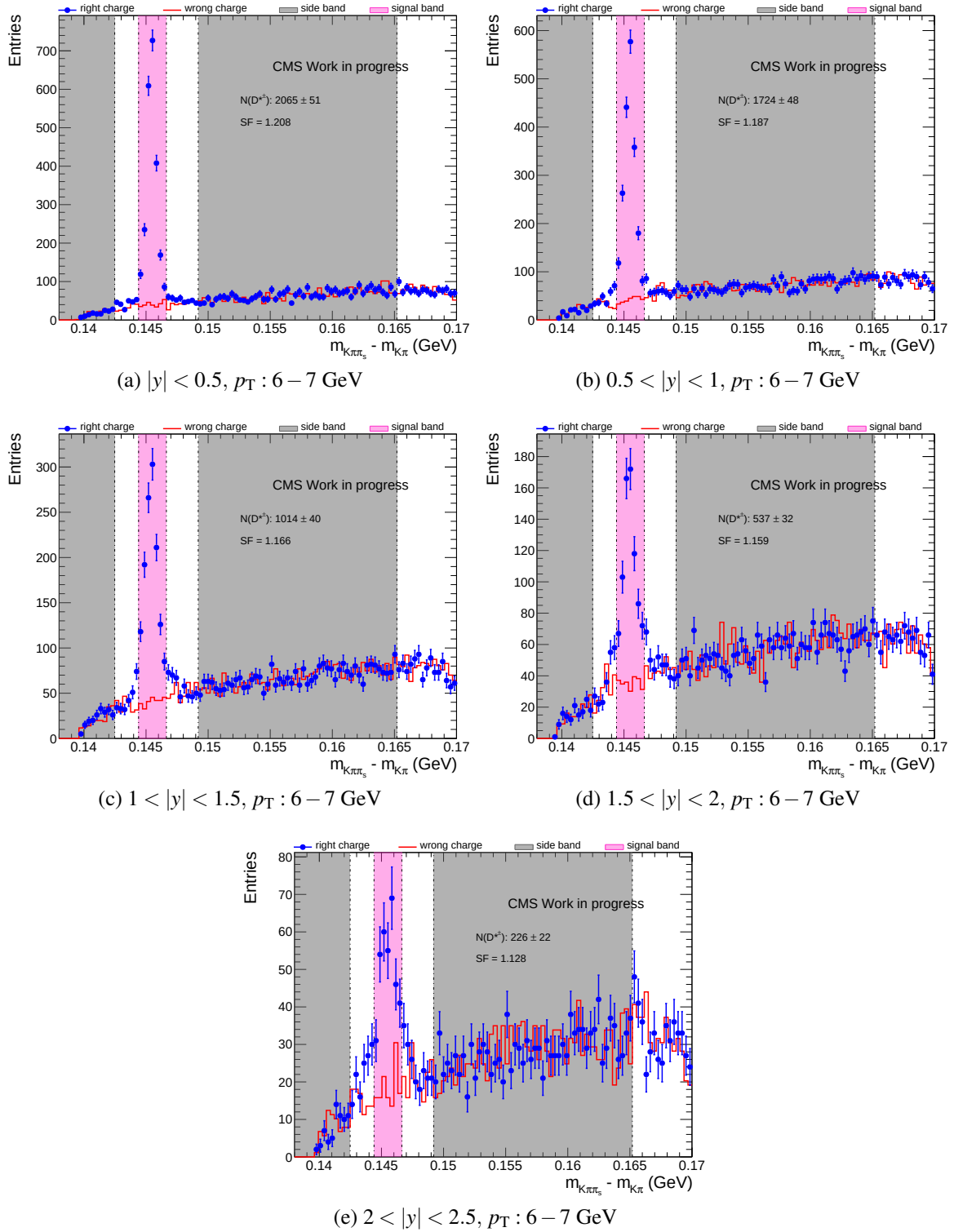
Fig. B.1 Signal extraction for $p_T : 1 - 2$ GeV and $0 < |y| < 2.5$

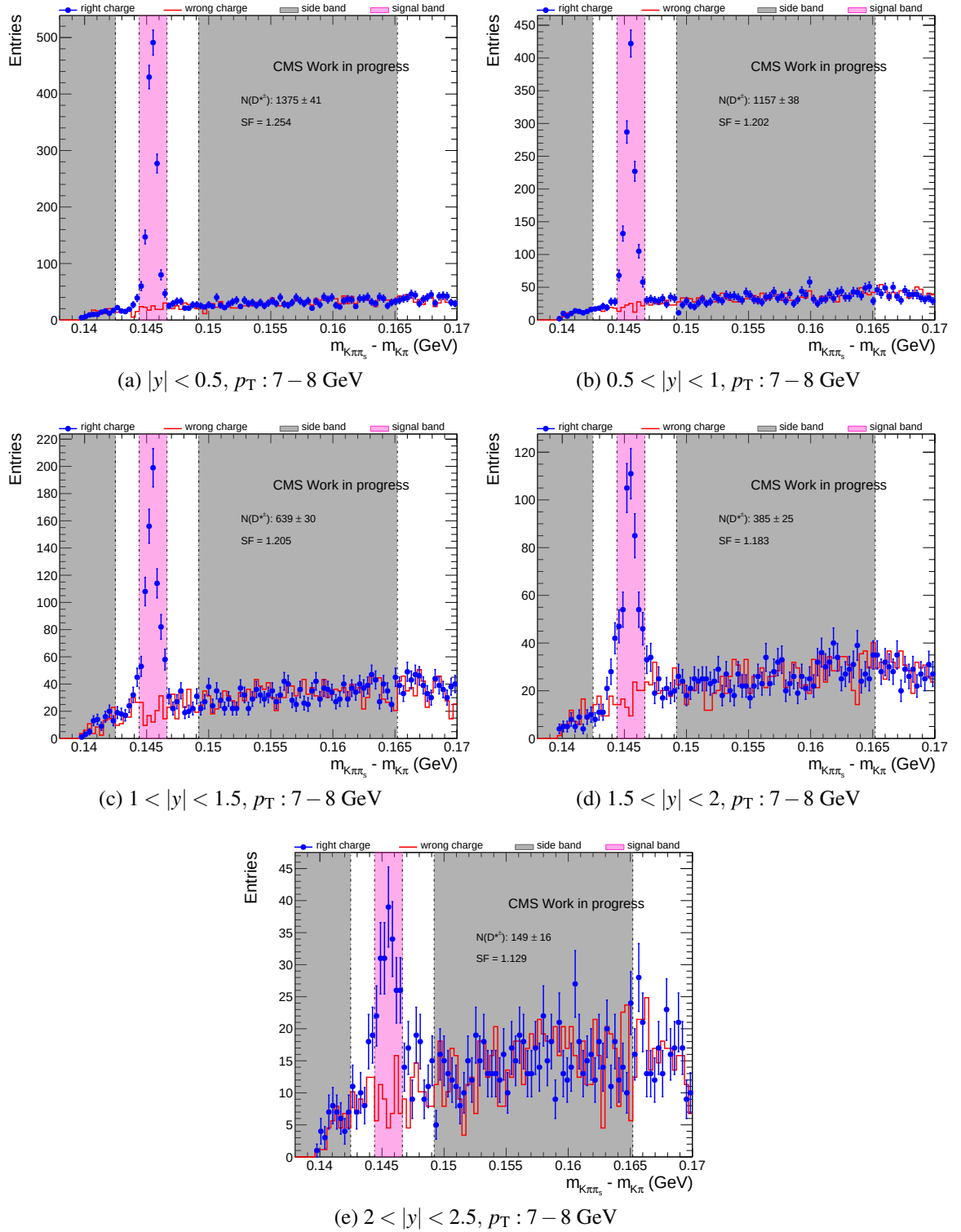
Fig. B.2 Signal extraction for $p_T : 2 - 3$ GeV and $0 < |y| < 2.5$

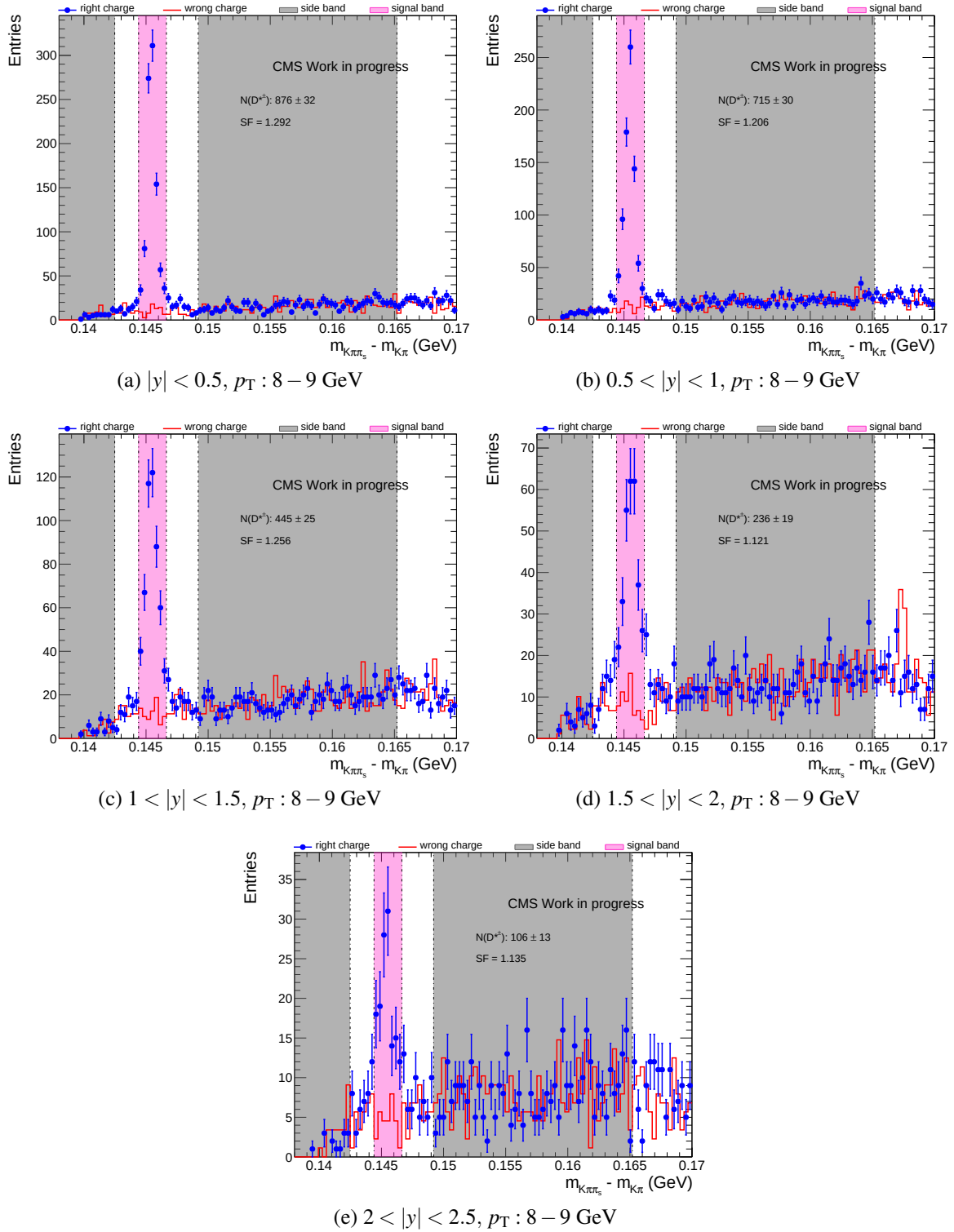
Fig. B.3 Signal extraction for $p_T : 3 - 4$ GeV and $0 < |y| < 2.5$

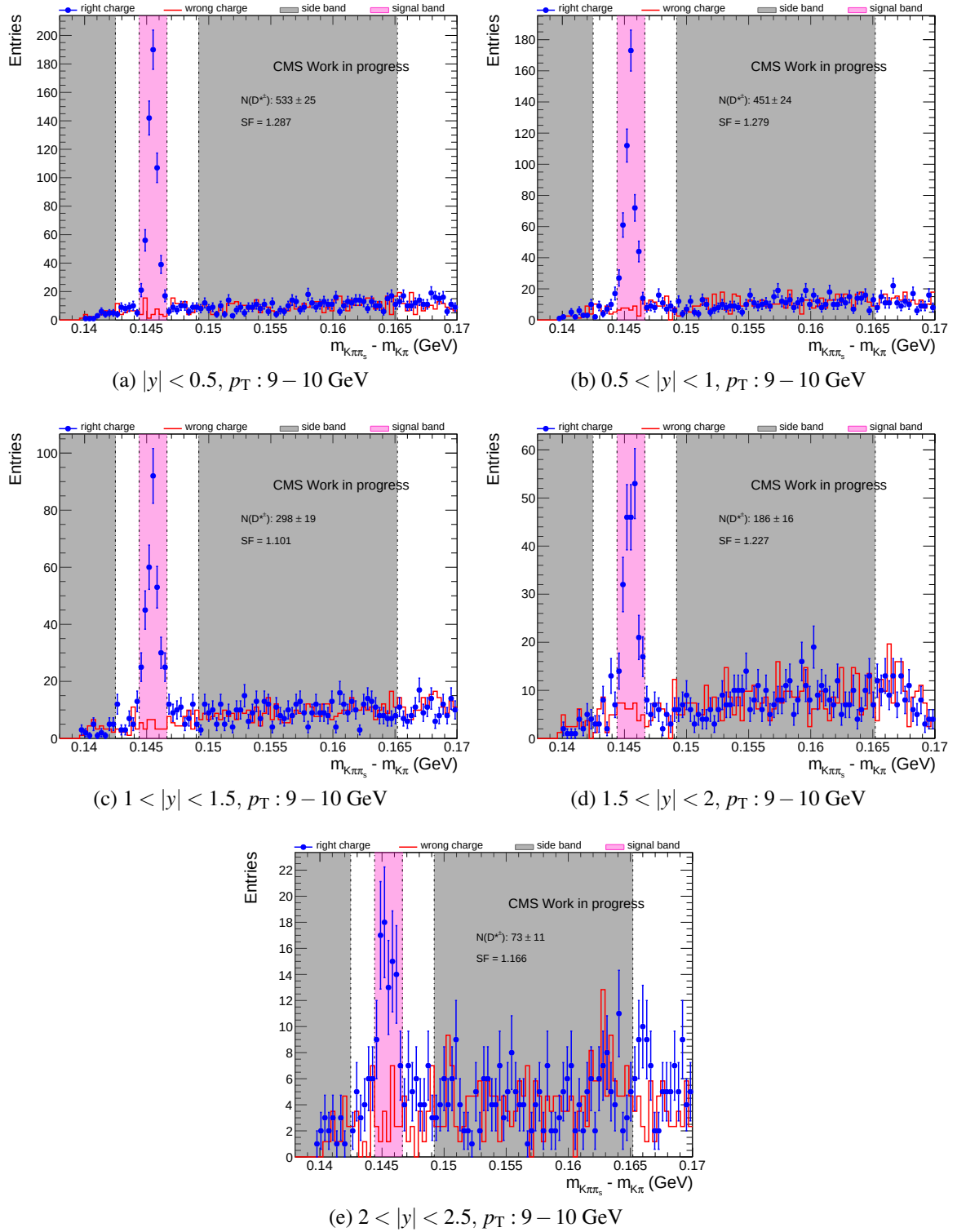
Fig. B.4 Signal extraction for $p_T : 4 - 5$ GeV and $0 < |y| < 2.5$

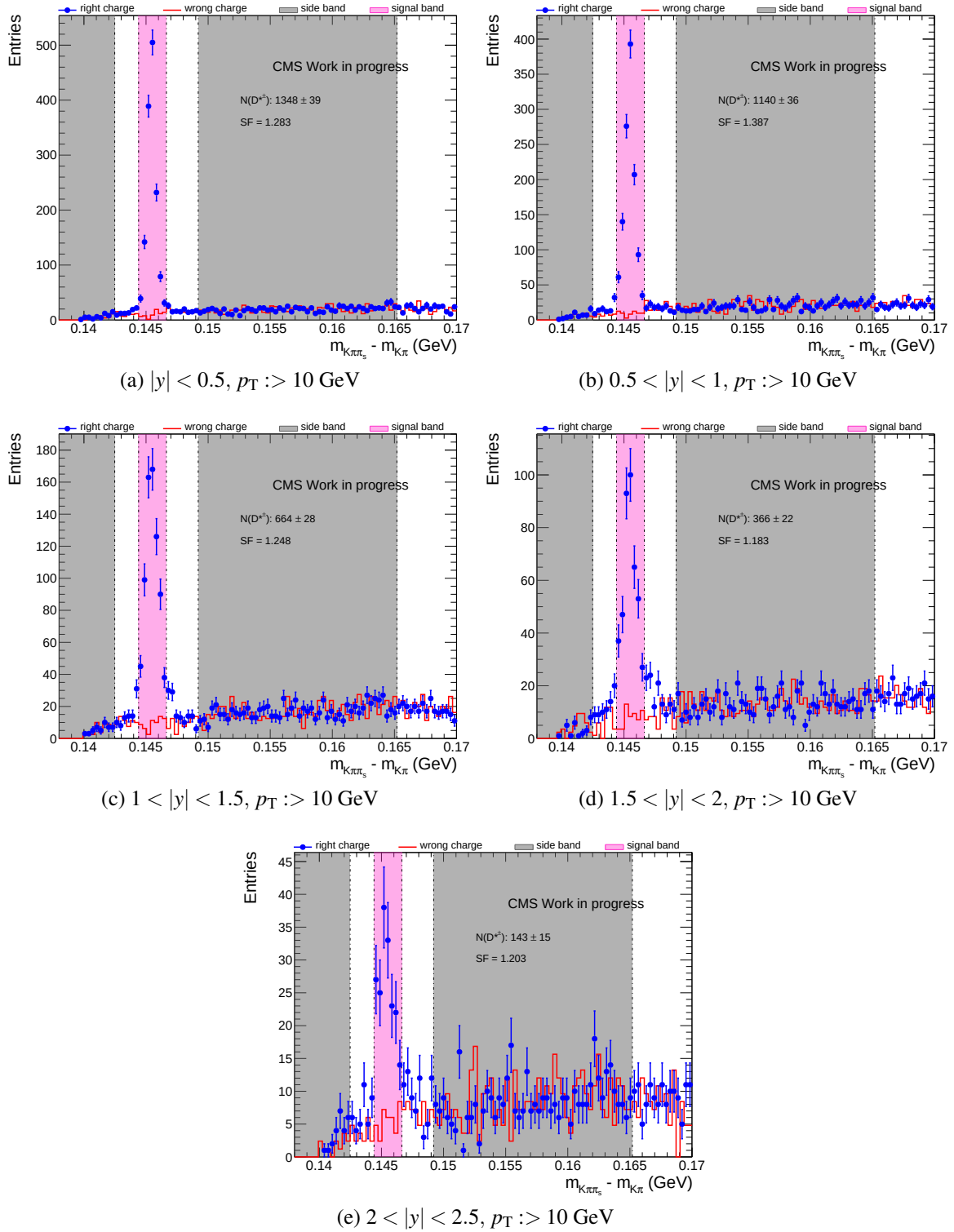
Fig. B.5 Signal extraction for $p_T : 5 - 6$ GeV and $0 < |y| < 2.5$

Fig. B.6 Signal extraction for $p_T : 6 - 7$ GeV and $0 < |y| < 2.5$

Fig. B.7 Signal extraction for $p_T : 7 - 8$ GeV and $0 < |y| < 2.5$

Fig. B.8 Signal extraction for $p_T : 8 - 9$ GeV and $0 < |y| < 2.5$

Fig. B.9 Signal extraction for $p_T : 9 - 10 \text{ GeV}$ and $0 < |y| < 2.5$

Fig. B.10 Signal extraction for $p_T :> 10 \text{ GeV}$ and $0 < |y| < 2.5$

References

- [1] CMS Collaboration, “CMS Luminosity Calibration for the pp Reference Run at $\sqrt{s} = 5.02$ TeV”,.

Appendix C

Measurement of Prompt and Non-prompt D^* Cross Sections in Bins of p_T and Rapidity

This Appendix is used to accumulate all ingredients which were necessary for the determination of the differential prompt and non-prompt D^* cross sections shown in Section 6.8. Due to statistical reason and for the purpose of the charm beauty separation explained in Section 6.6 the rapidity region is divided into $|y| < 0.5$, $|y| < 1$ and $|y| < 2$. The signal extraction for D^* for these bins of p_T and rapidity are shown in Section C.1. For the separation of prompt and non-prompt D^* the signal extraction at truth level were needed for prompt and non-prompt D^* which are listed in Section C.2. The charm beauty separation in D^* final states is for these bins in the phase space performed. The method via dca and via $dca \cdot dl_{sig}$ is shown in Section C.3. Correlation plots to investigate potential correlations between observables are shown in Section C.4.

C.1 Signal Extraction for D^* in Bins of p_T and Rapidity

This sections shows the signal extraction for D^* in bins bins of p_T and rapidity.

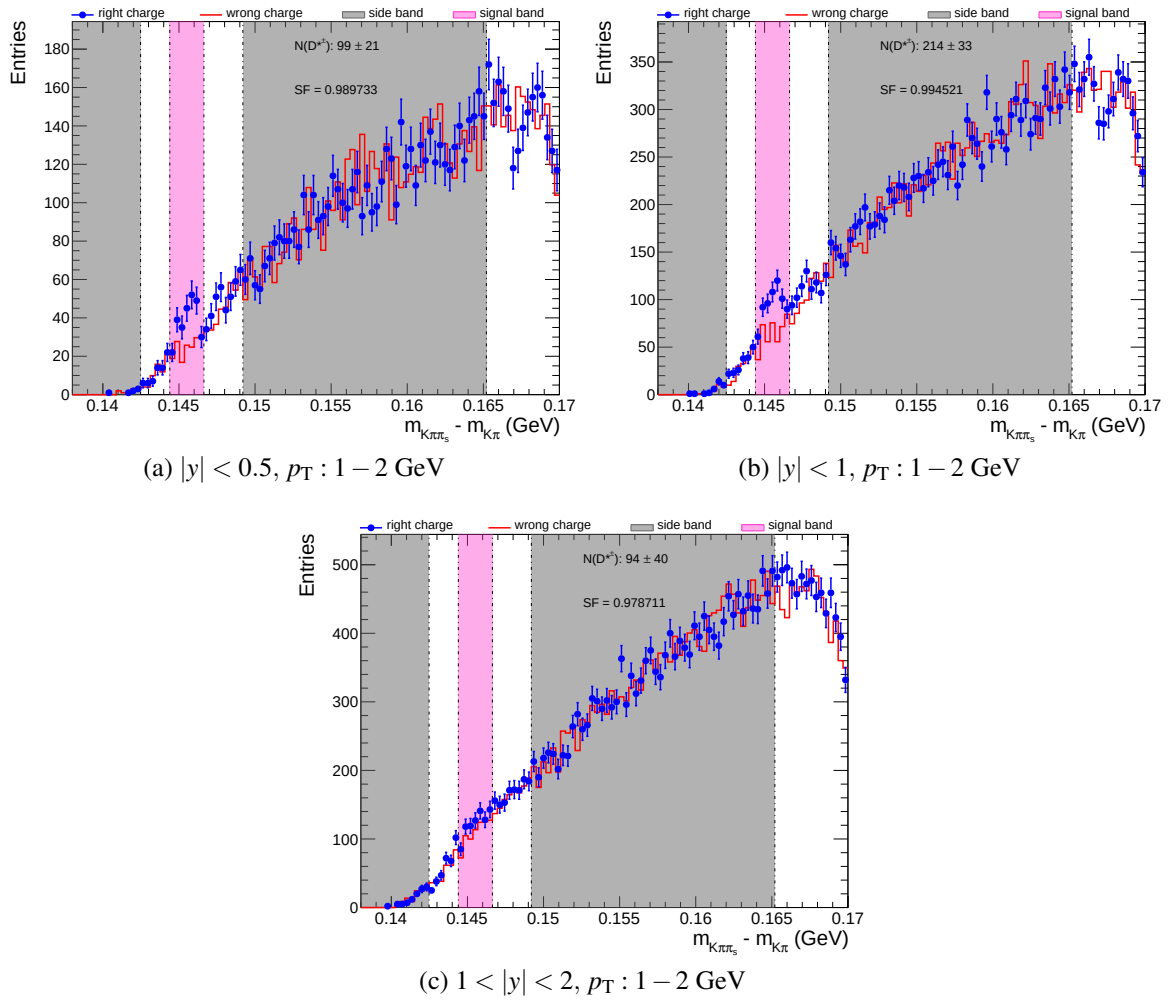
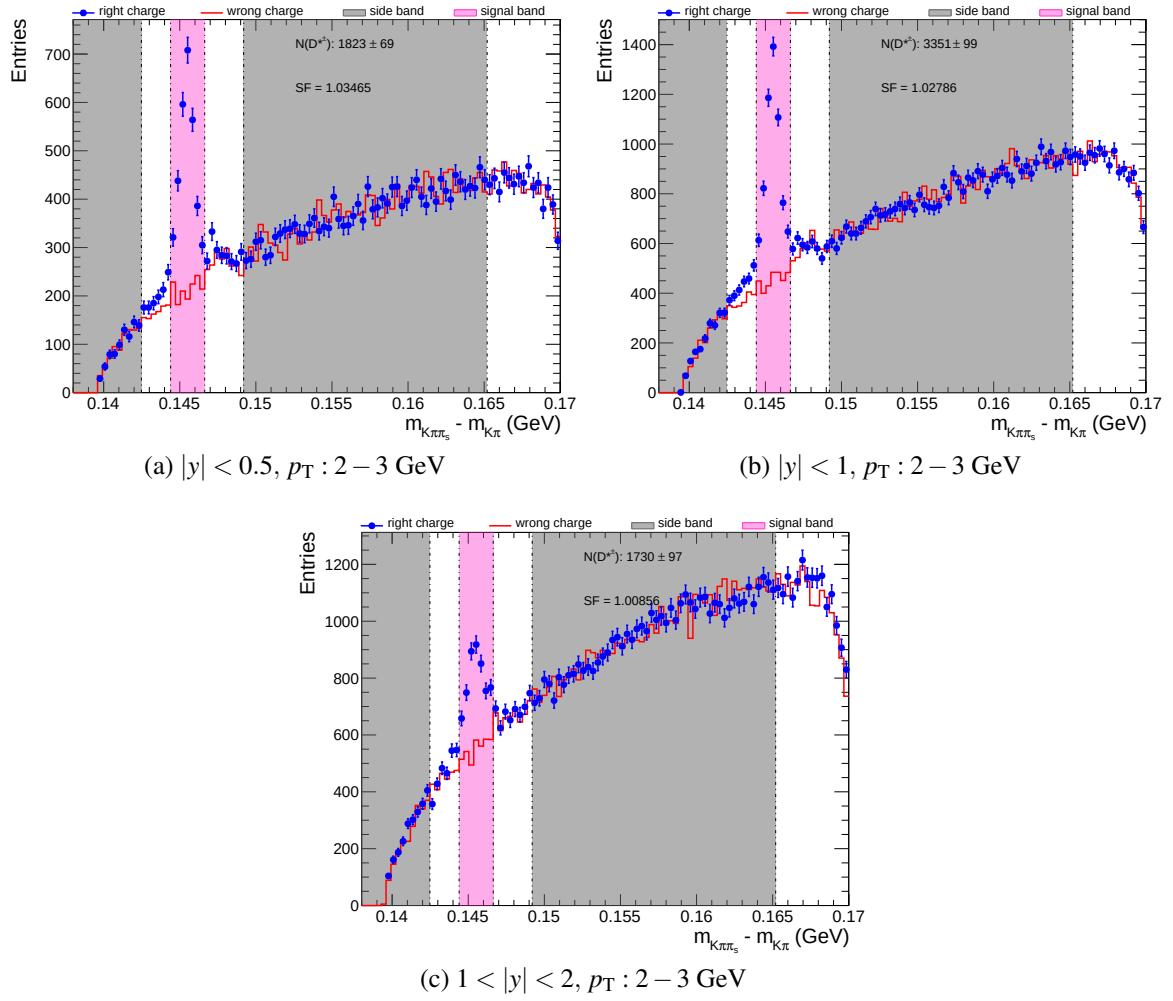


Fig. C.1 Signal extraction for $p_T : 1 - 2$ GeV and $0 < |y| < 2$

Fig. C.2 Signal extraction for $p_T : 2 - 3$ GeV and $0 < |y| < 2$

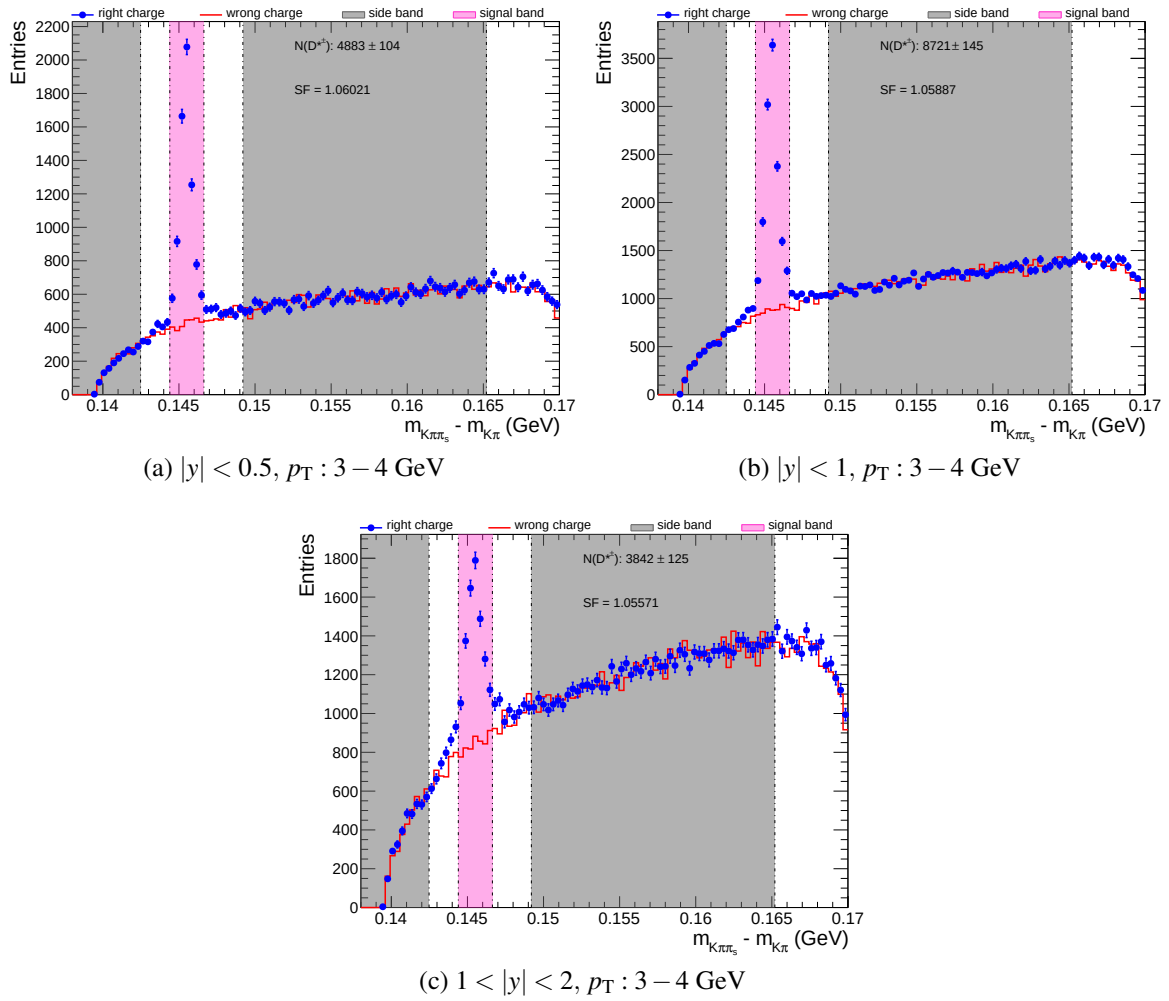
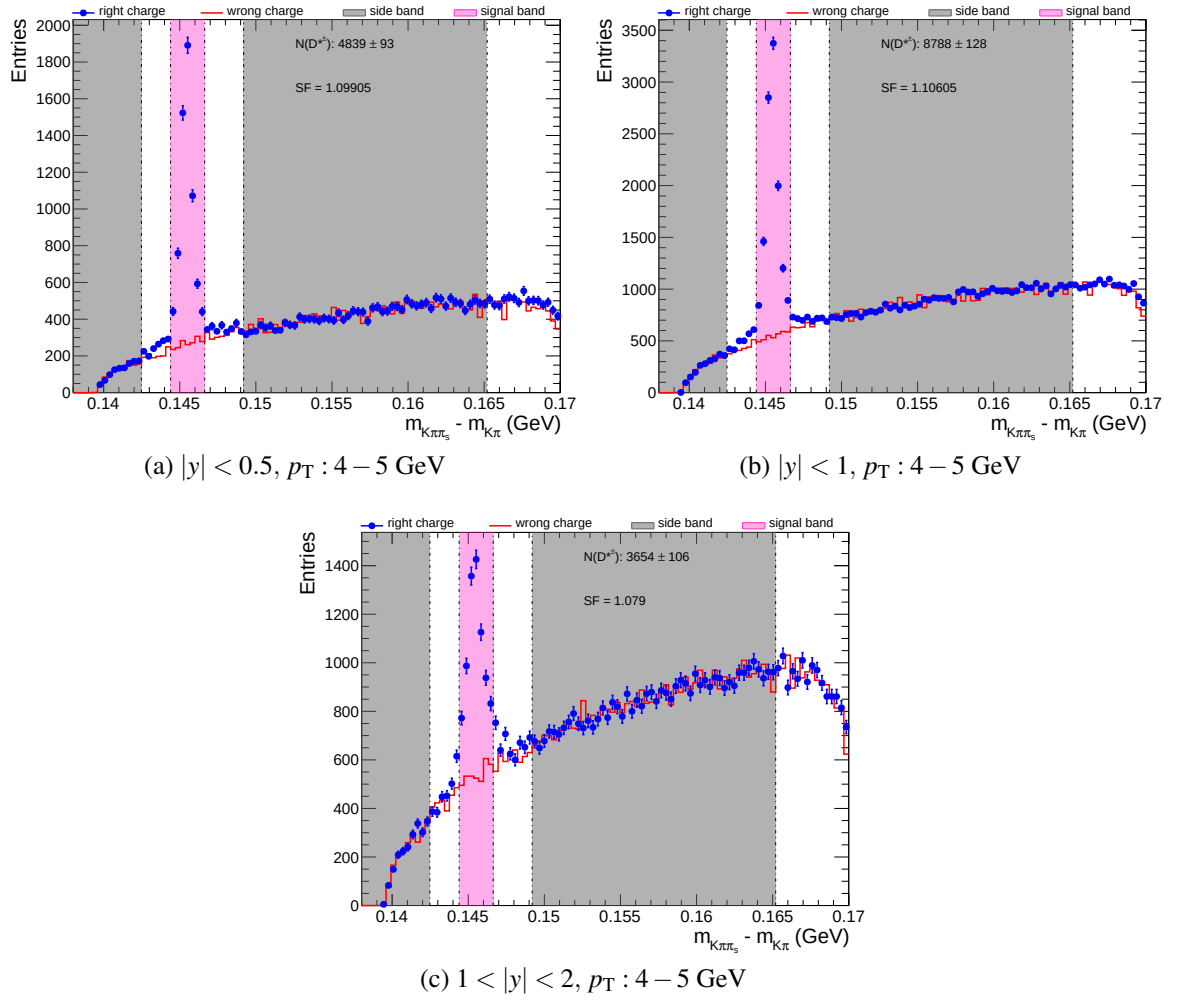


Fig. C.3 Signal extraction for $p_T : 3 - 4$ GeV and $0 < |y| < 2$

Fig. C.4 Signal extraction for $p_T : 4 - 5$ GeV and $0 < |y| < 2$

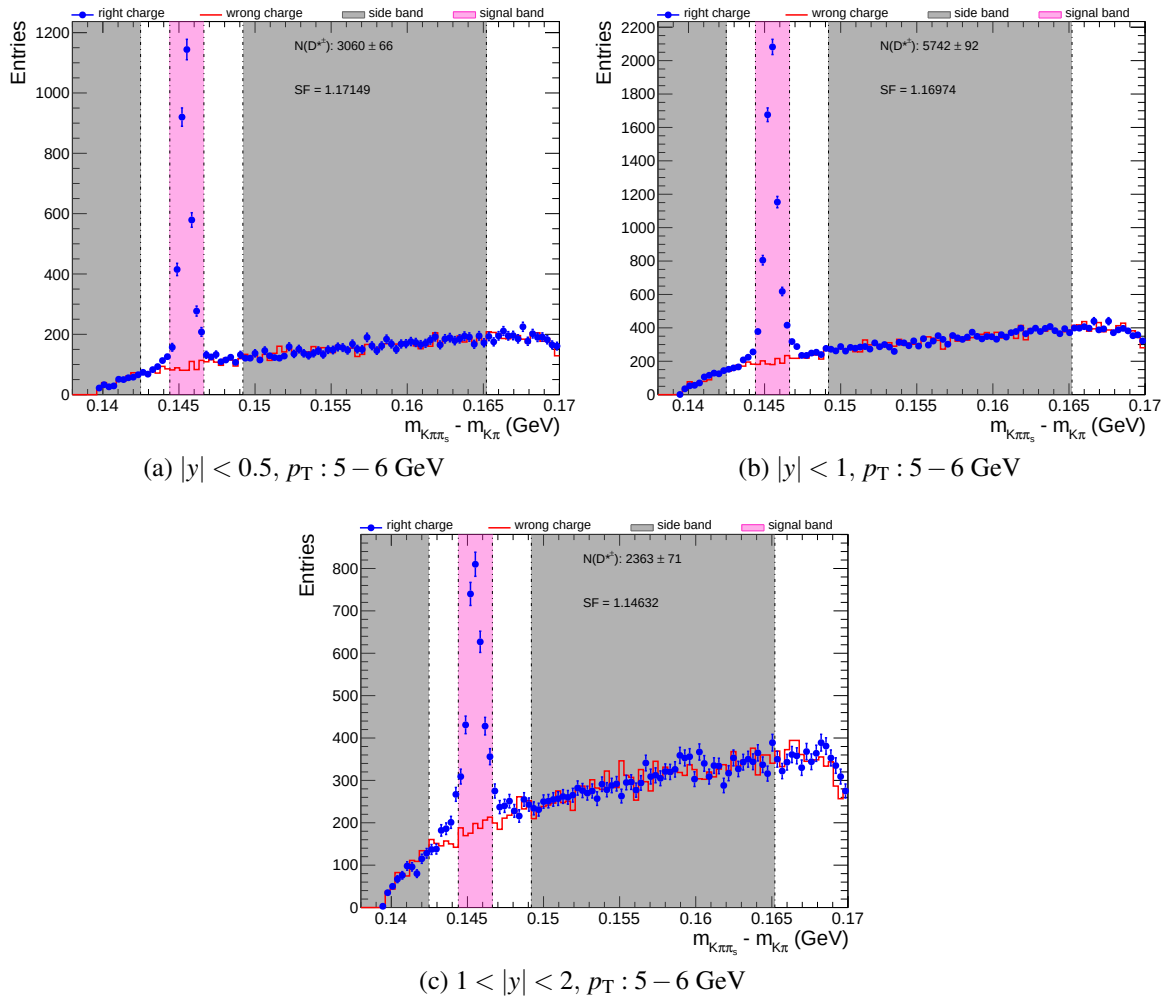
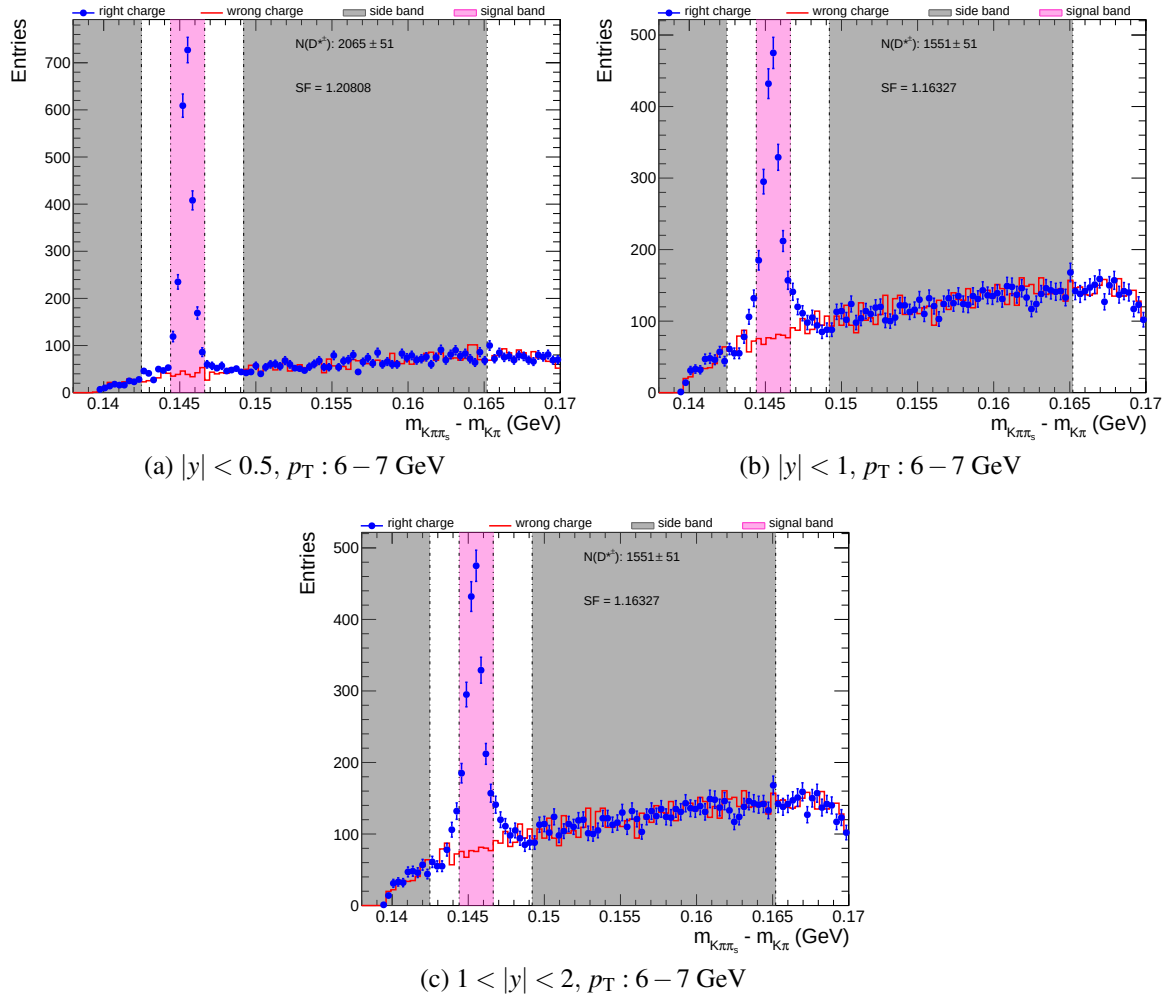


Fig. C.5 Signal extraction for $p_T : 5 - 6$ GeV and $0 < |y| < 2$

Fig. C.6 Signal extraction for $p_T : 6 - 7$ GeV and $0 < |y| < 2$

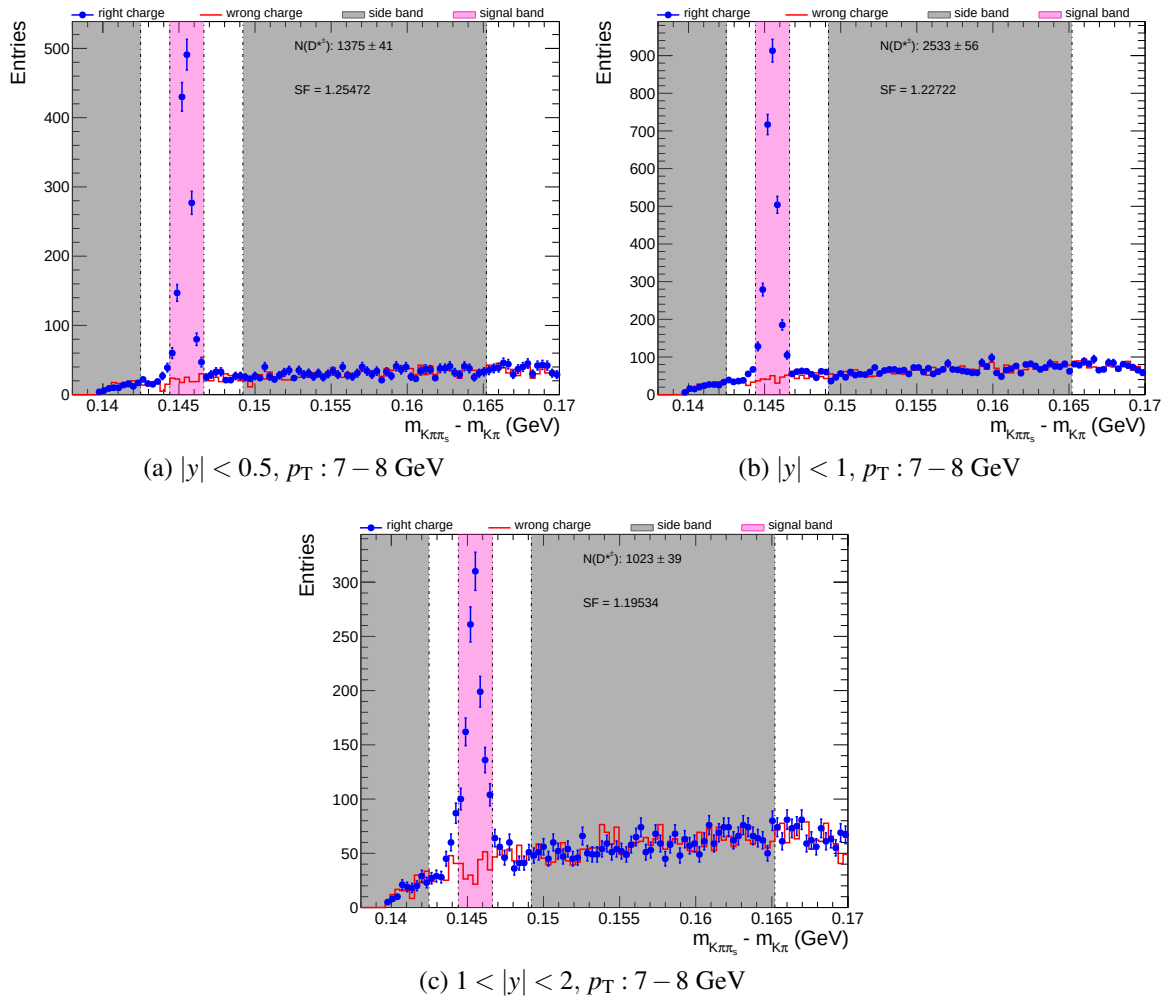
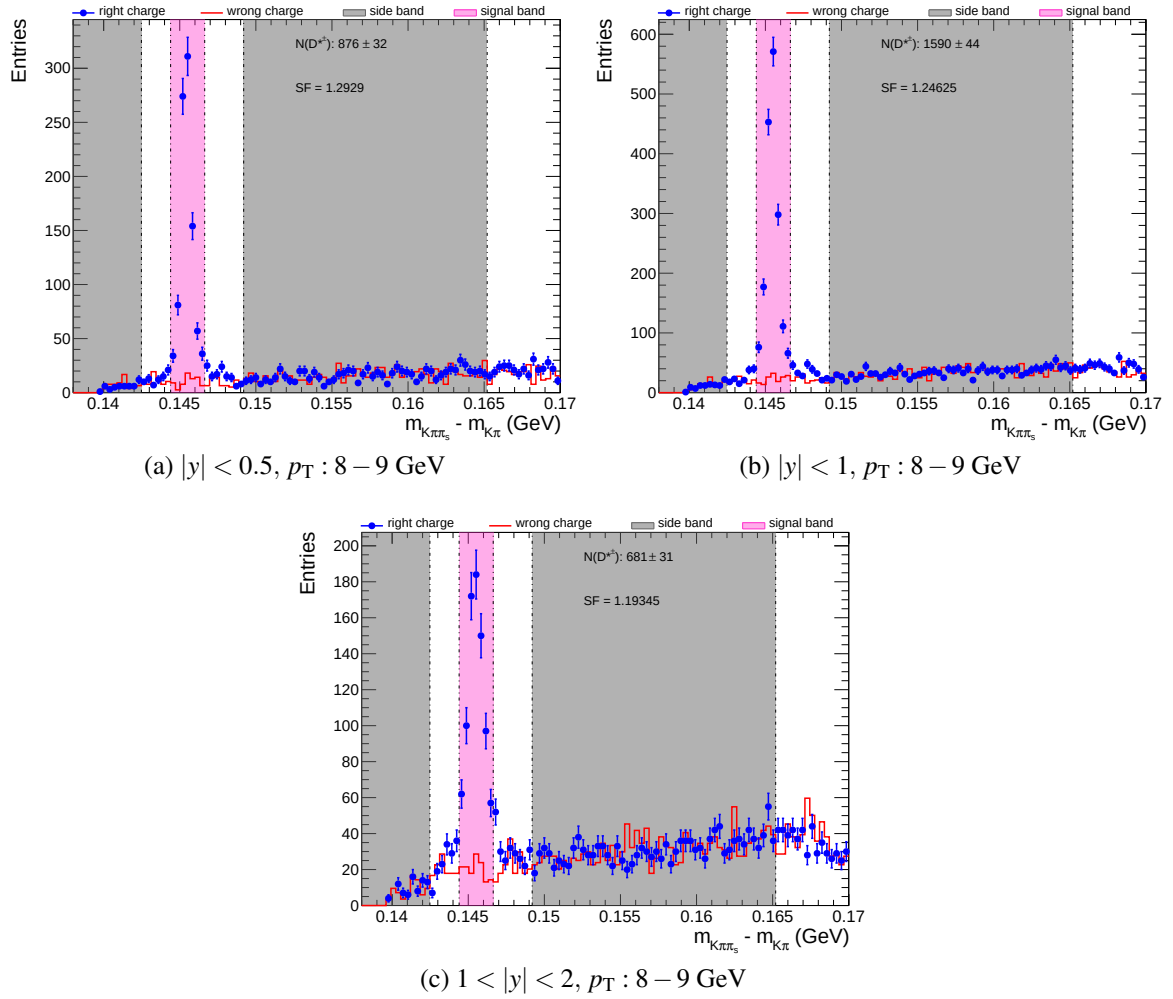


Fig. C.7 Signal extraction for $p_T : 7 - 8 \text{ GeV}$ and $0 < |y| < 2$

Fig. C.8 Signal extraction for $p_T : 8 - 9$ GeV and $0 < |y| < 2$

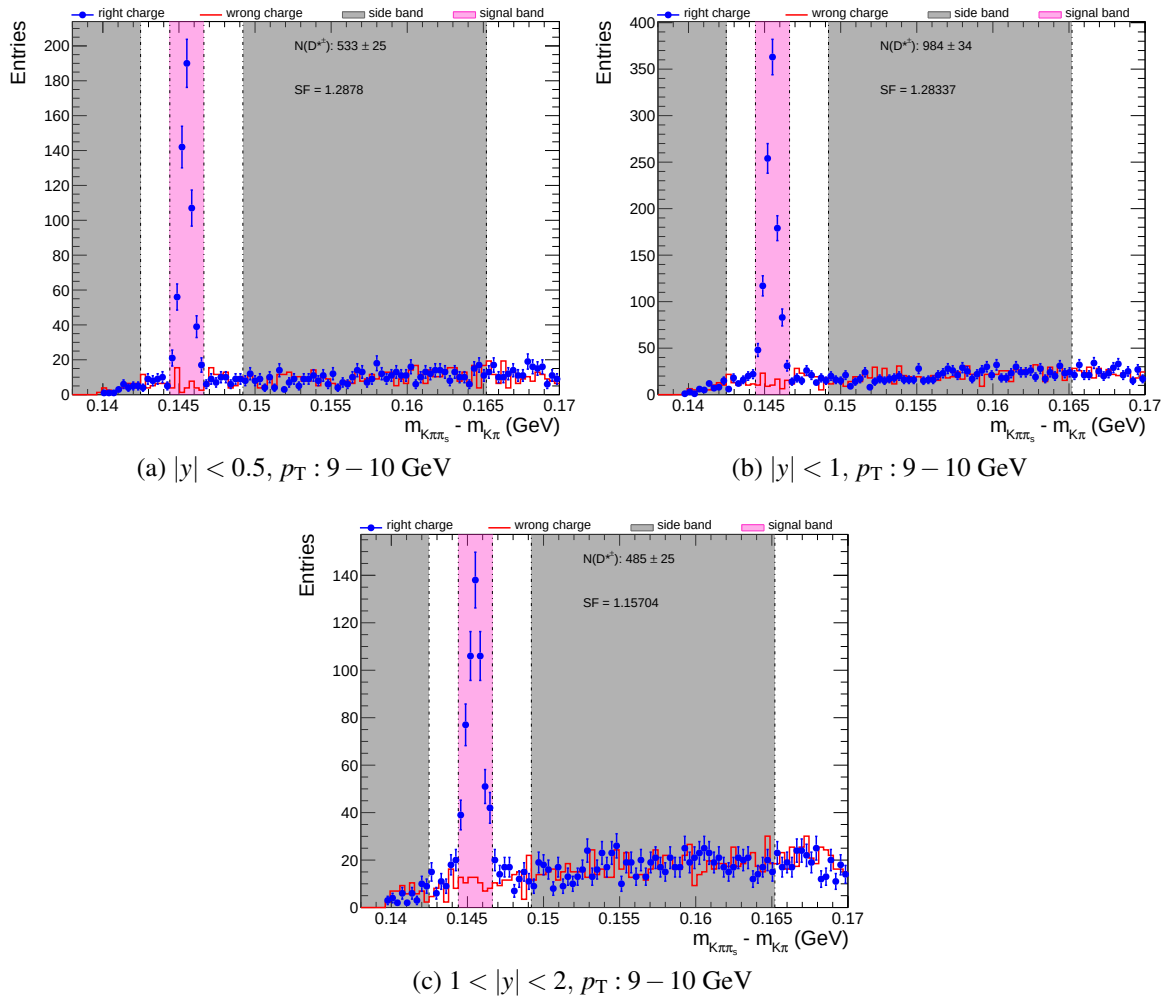
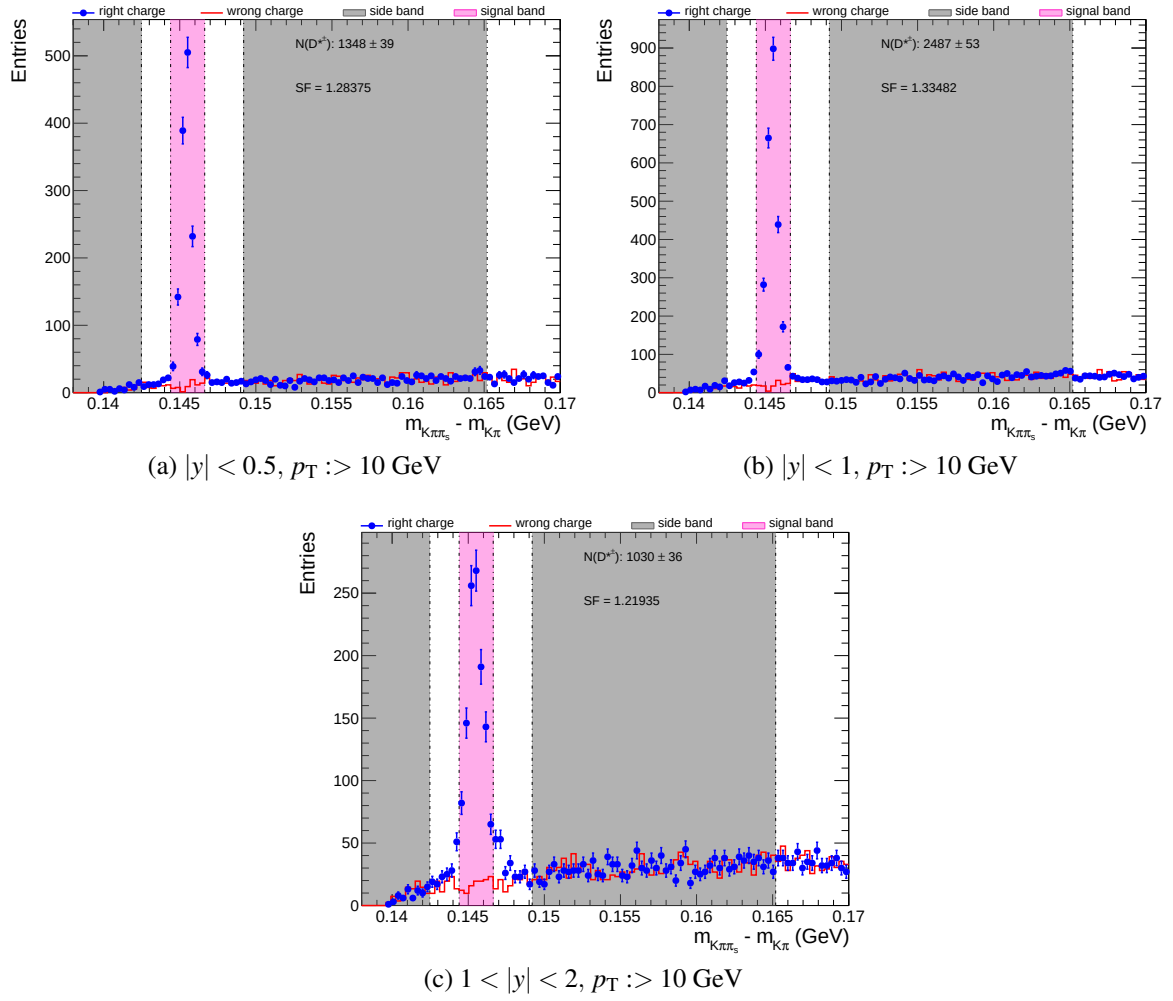


Fig. C.9 Signal extraction for $p_T : 9 - 10$ GeV and $0 < |y| < 2$

Fig. C.10 Signal extraction for $p_T :> 10 \text{ GeV}$ and $0 < |y| < 2$

C.2 Signal Extraction at Truth Level for Prompt and Non-prompt D^* in Bins of p_T and Rapidity

In the following the individual signal extraction for prompt D^* at truth level.

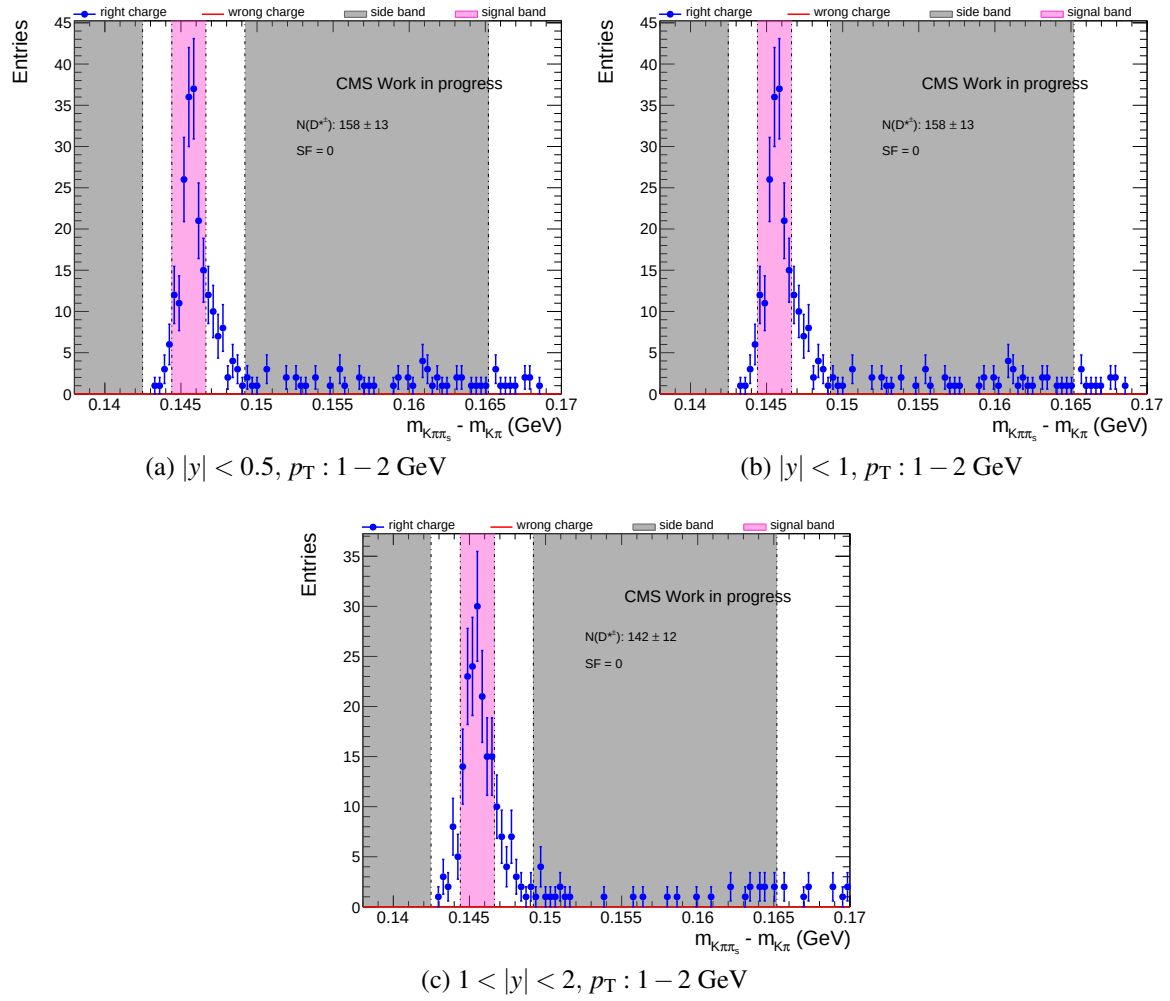


Fig. C.11 Signal extraction in MC for $p_T : 1 - 2$ GeV and $0 < |y| < 2$

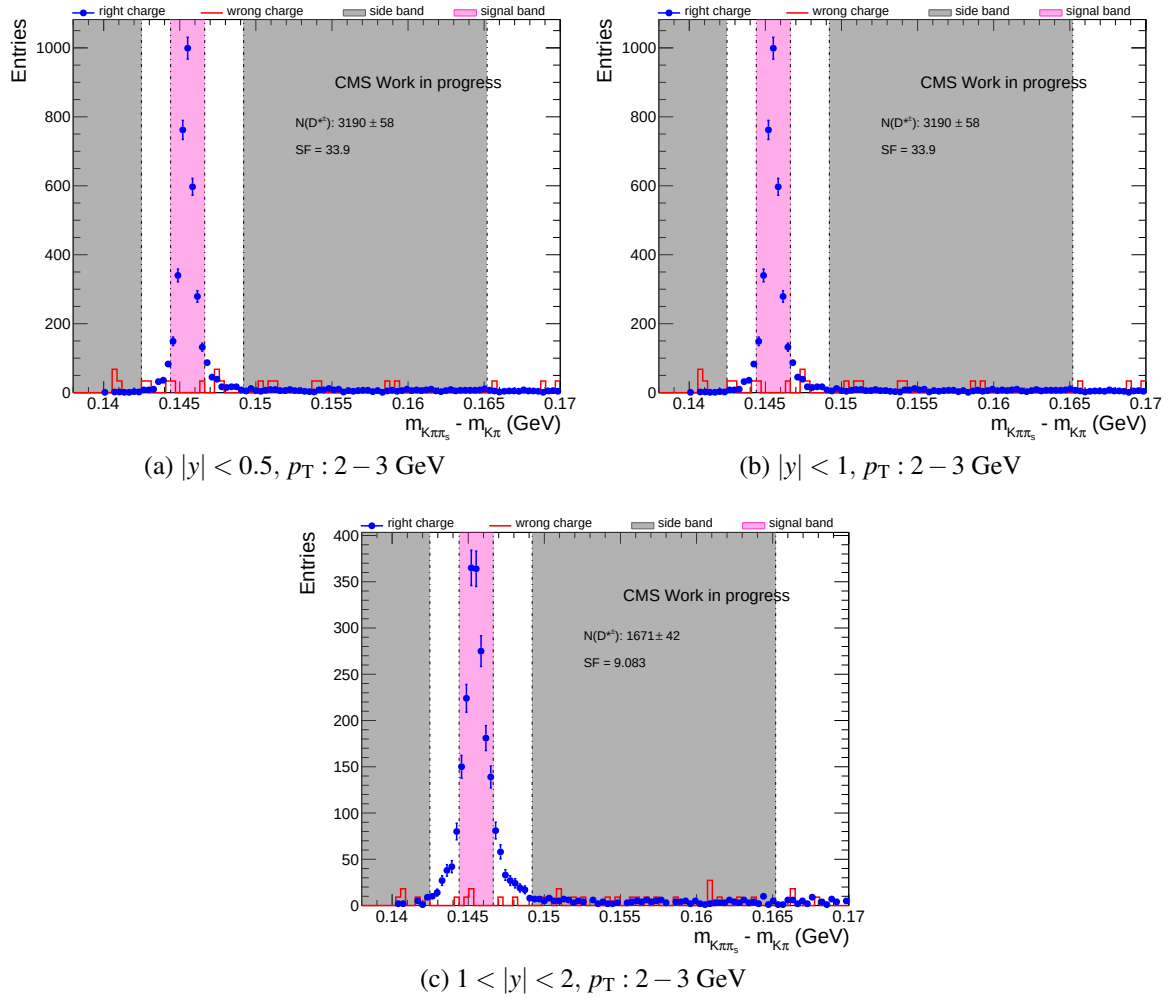


Fig. C.12 Signal extraction in MC for $p_T : 2 - 3$ GeV and $0 < |y| < 2$

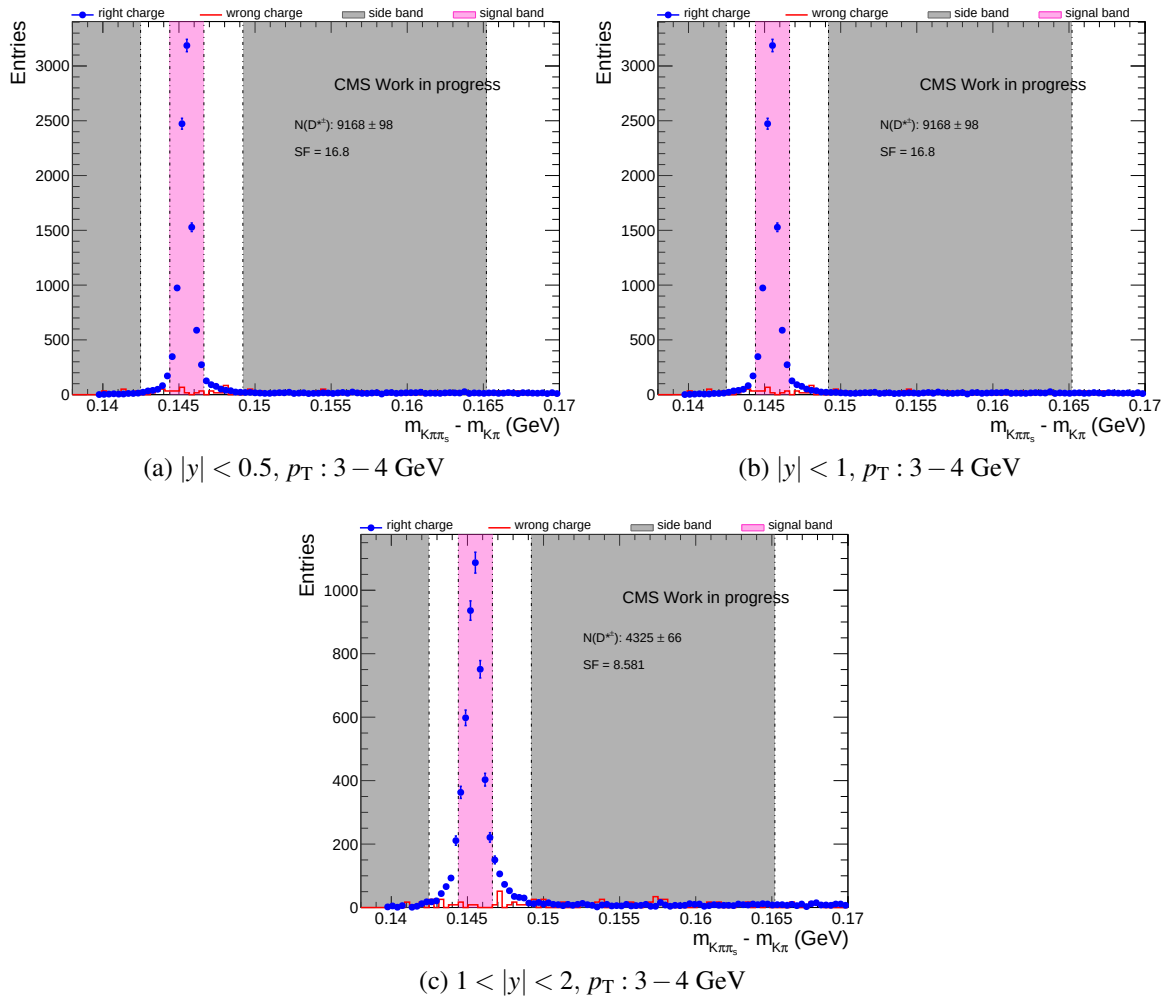


Fig. C.13 Signal extraction in MC for $p_T : 3 - 4$ GeV and $0 < |y| < 2$

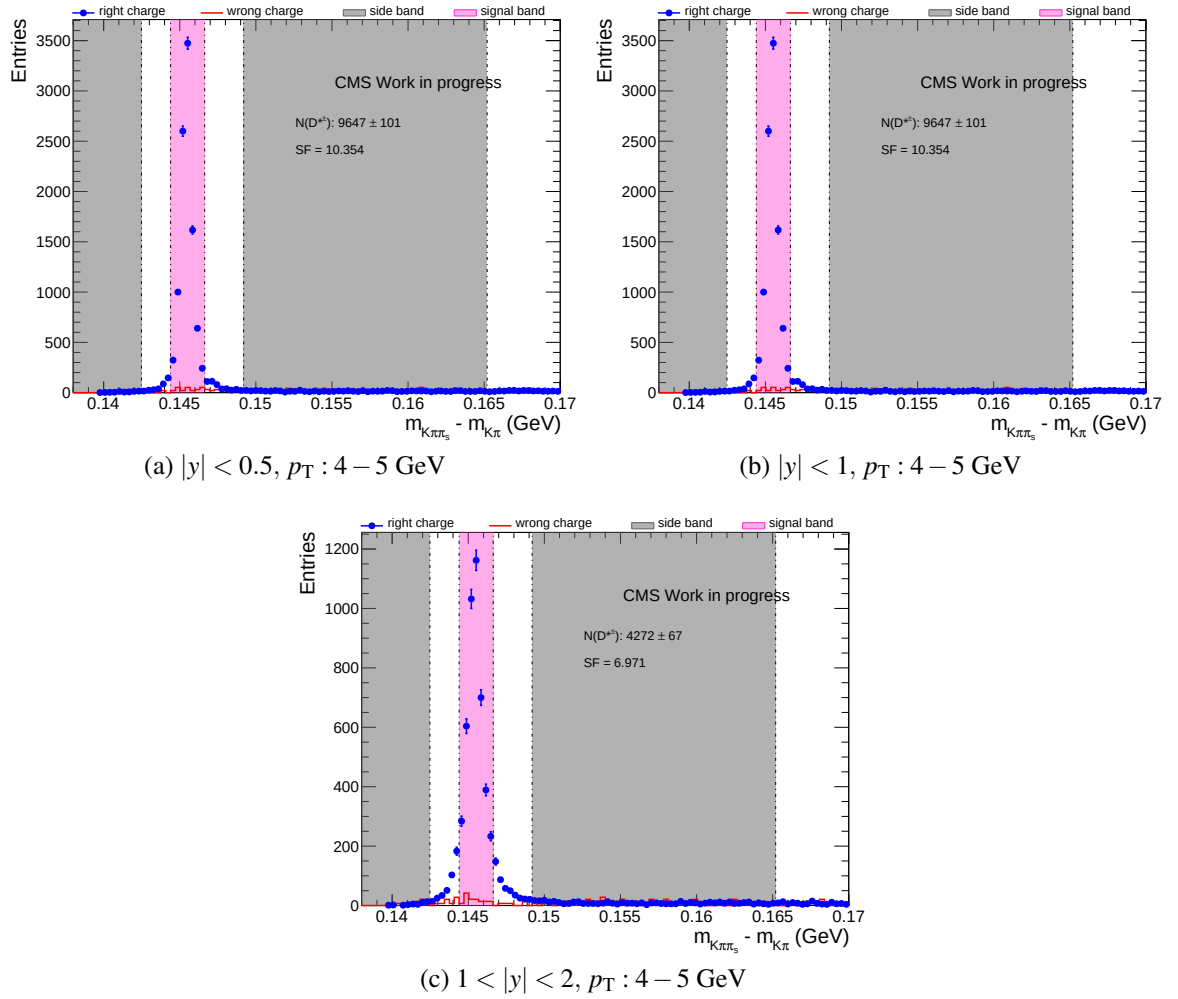


Fig. C.14 Signal extraction in MC for $p_T : 4 - 5$ GeV and $0 < |y| < 2$

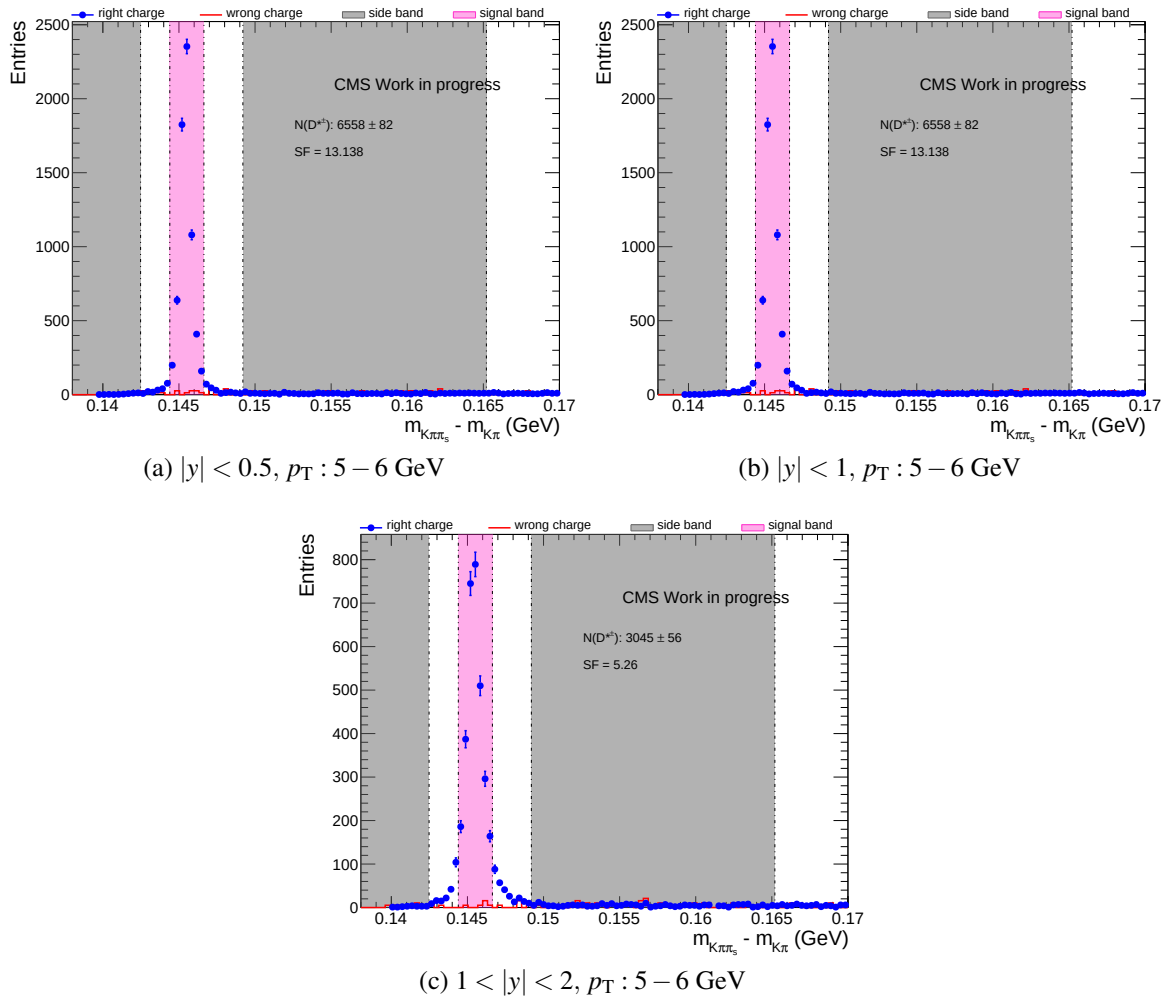


Fig. C.15 Signal extraction in MC for $p_T : 5 - 6$ GeV and $0 < |y| < 2$

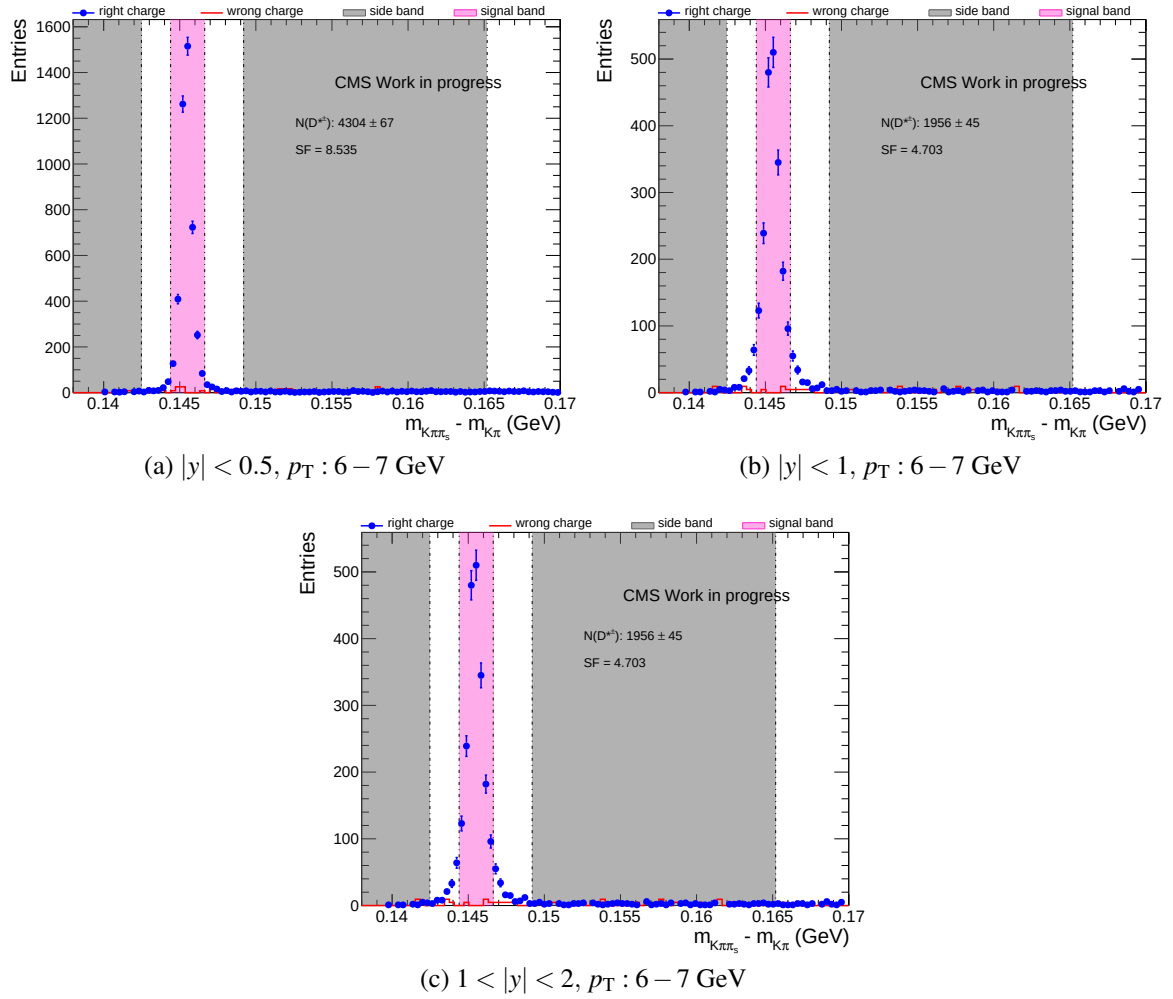


Fig. C.16 Signal extraction in MC for $p_T : 6 - 7$ GeV and $0 < |y| < 2$

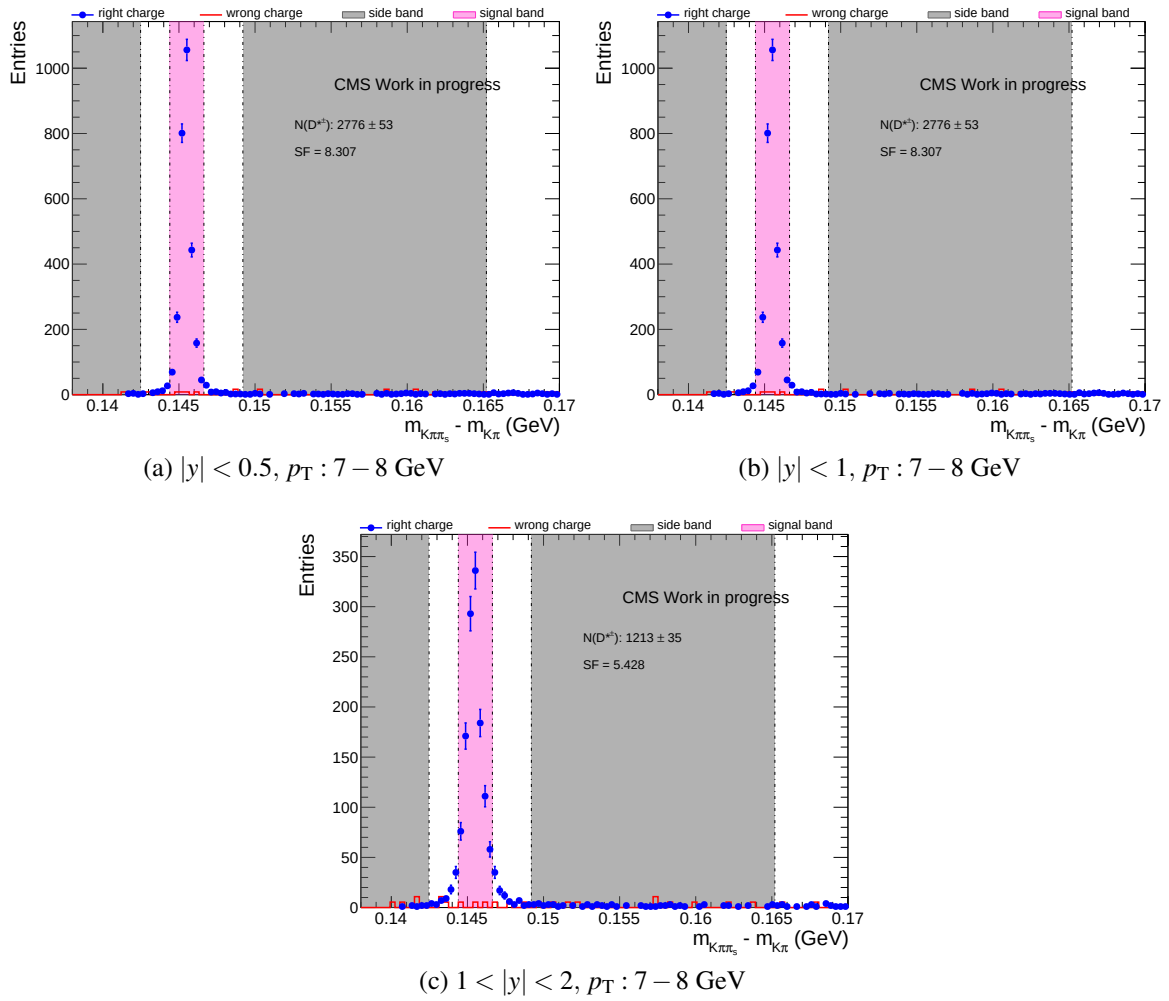


Fig. C.17 Signal extraction in MC for $p_T : 7 - 8 \text{ GeV}$ and $0 < |y| < 2$

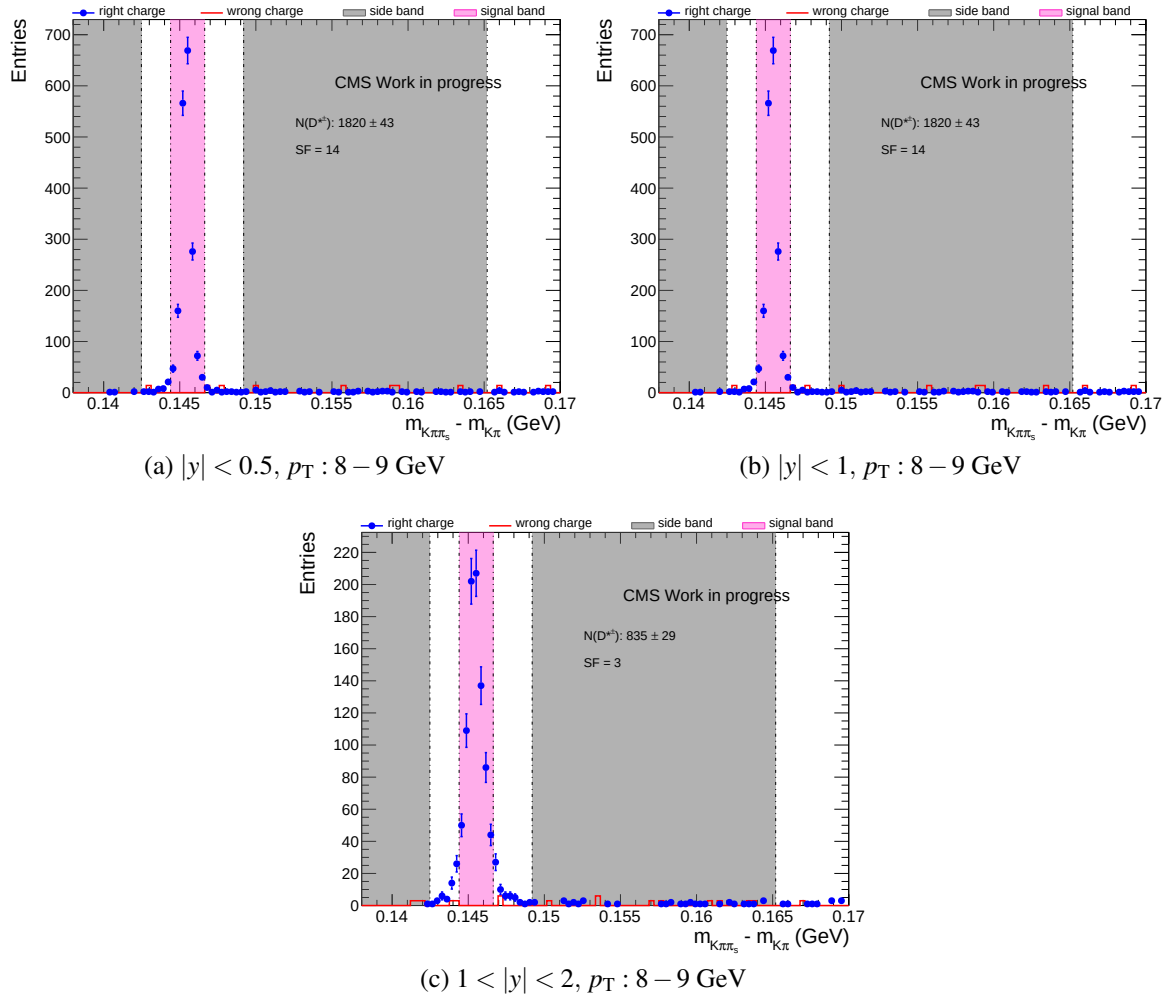


Fig. C.18 Signal extraction in MC for $p_T : 8 - 9$ GeV and $0 < |y| < 2$

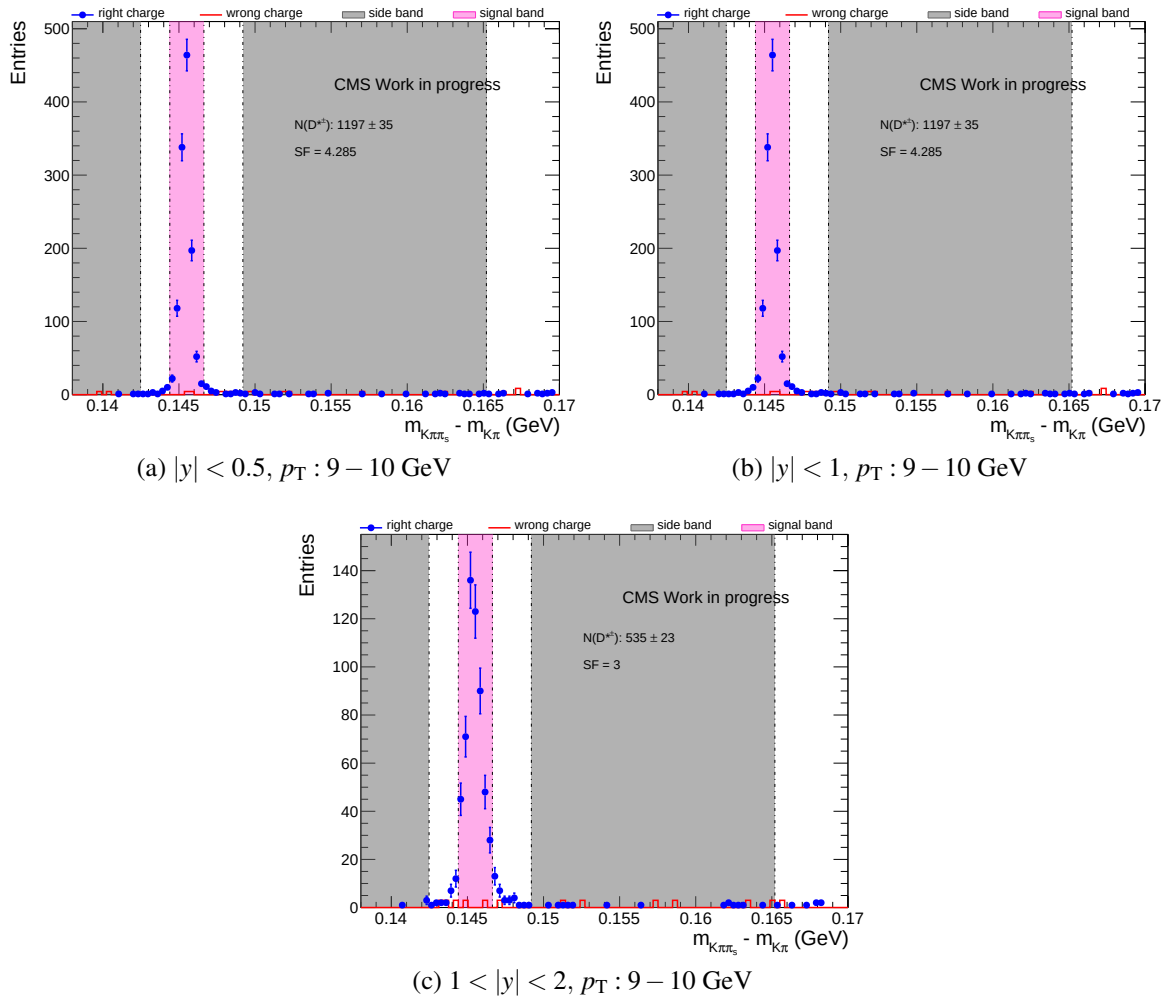


Fig. C.19 Signal extraction in MC for $p_T : 9 - 10$ GeV and $0 < |y| < 2$

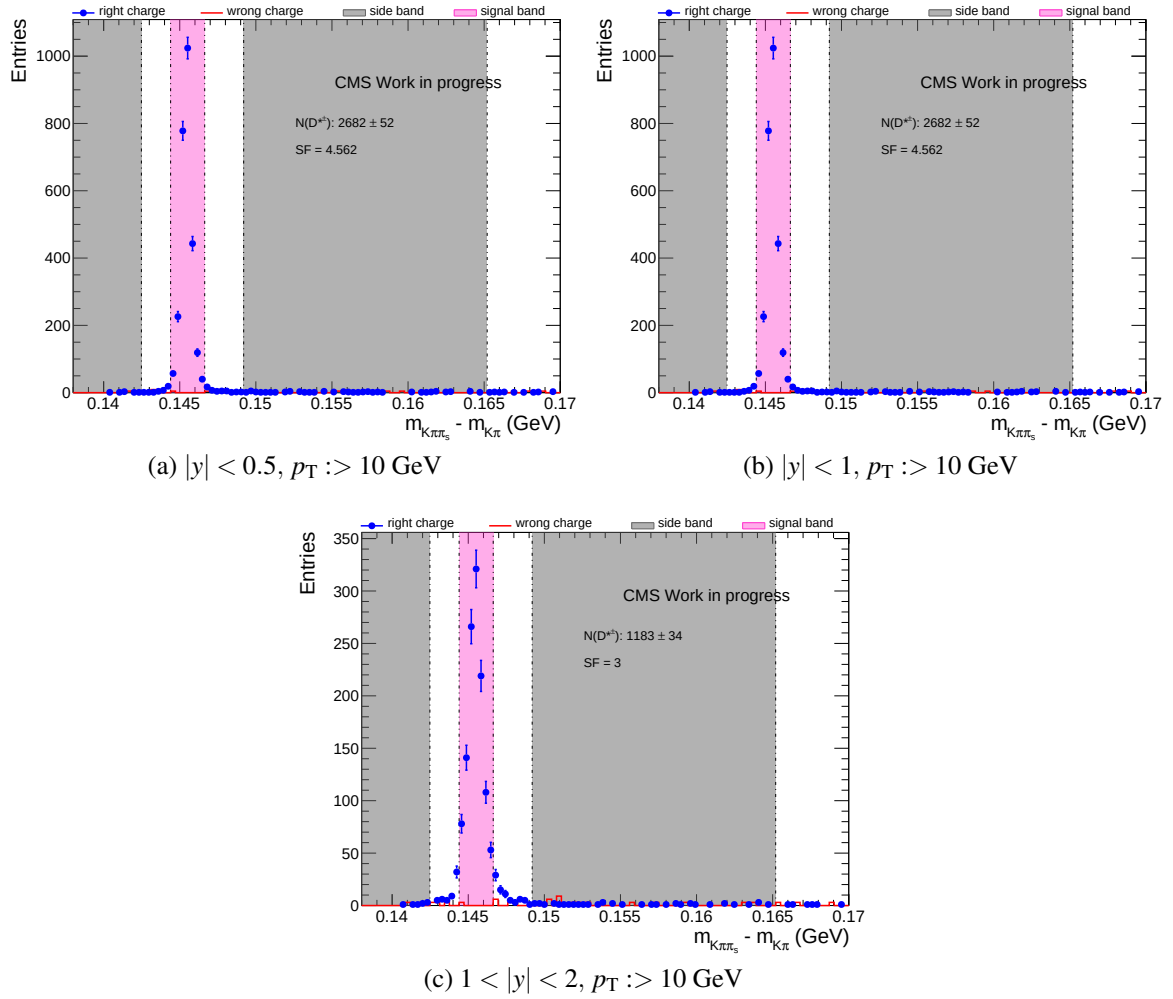


Fig. C.20 Signal extraction in MC for $p_T :> 10 \text{ GeV}$ and $0 < |y| < 2$

Now in the following the individual Signal extraction in MC for non-prompt D^* at truth level.

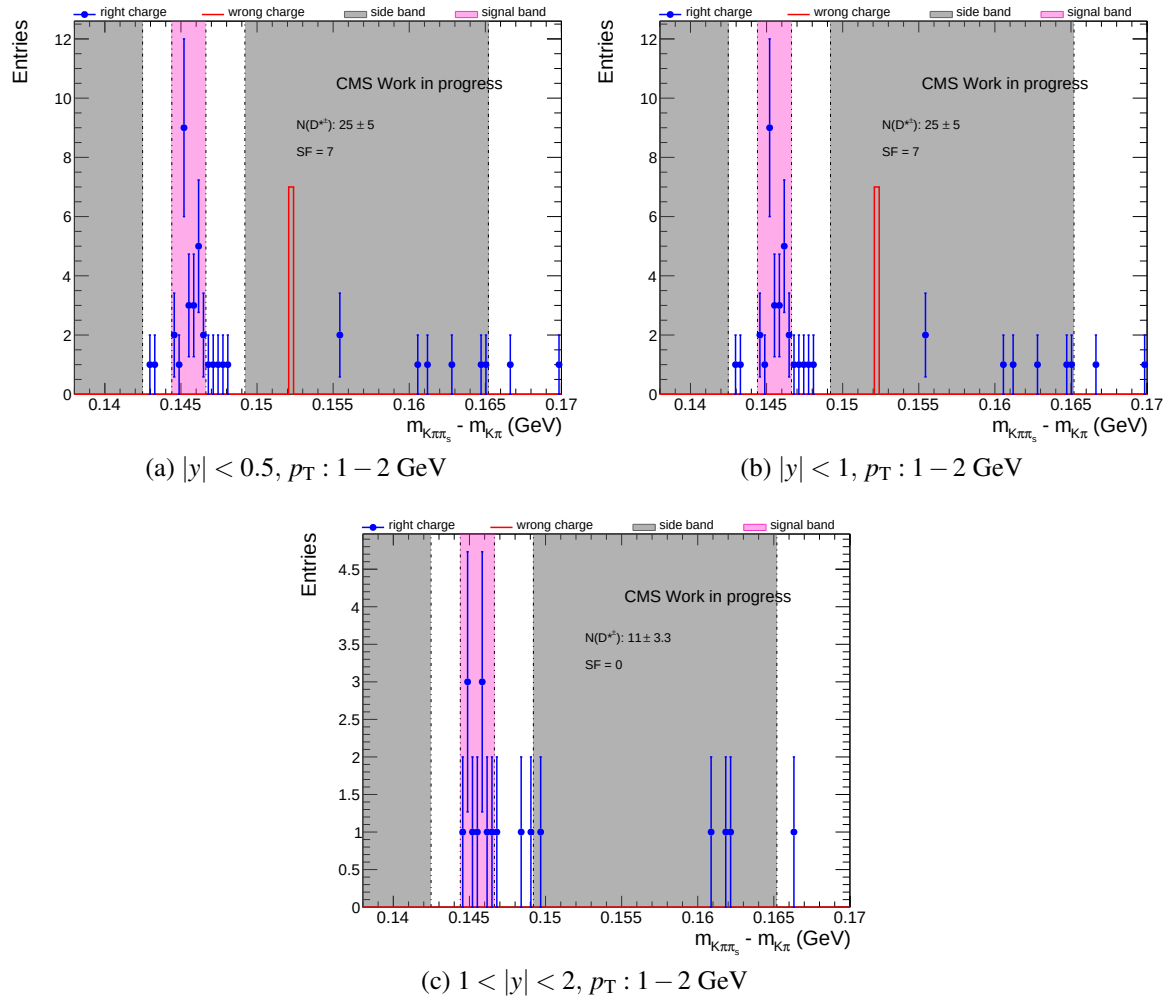
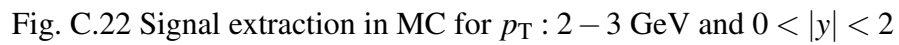


Fig. C.21 Signal extraction in MC for $p_T : 1 - 2 \text{ GeV}$ and $0 < |y| < 2$



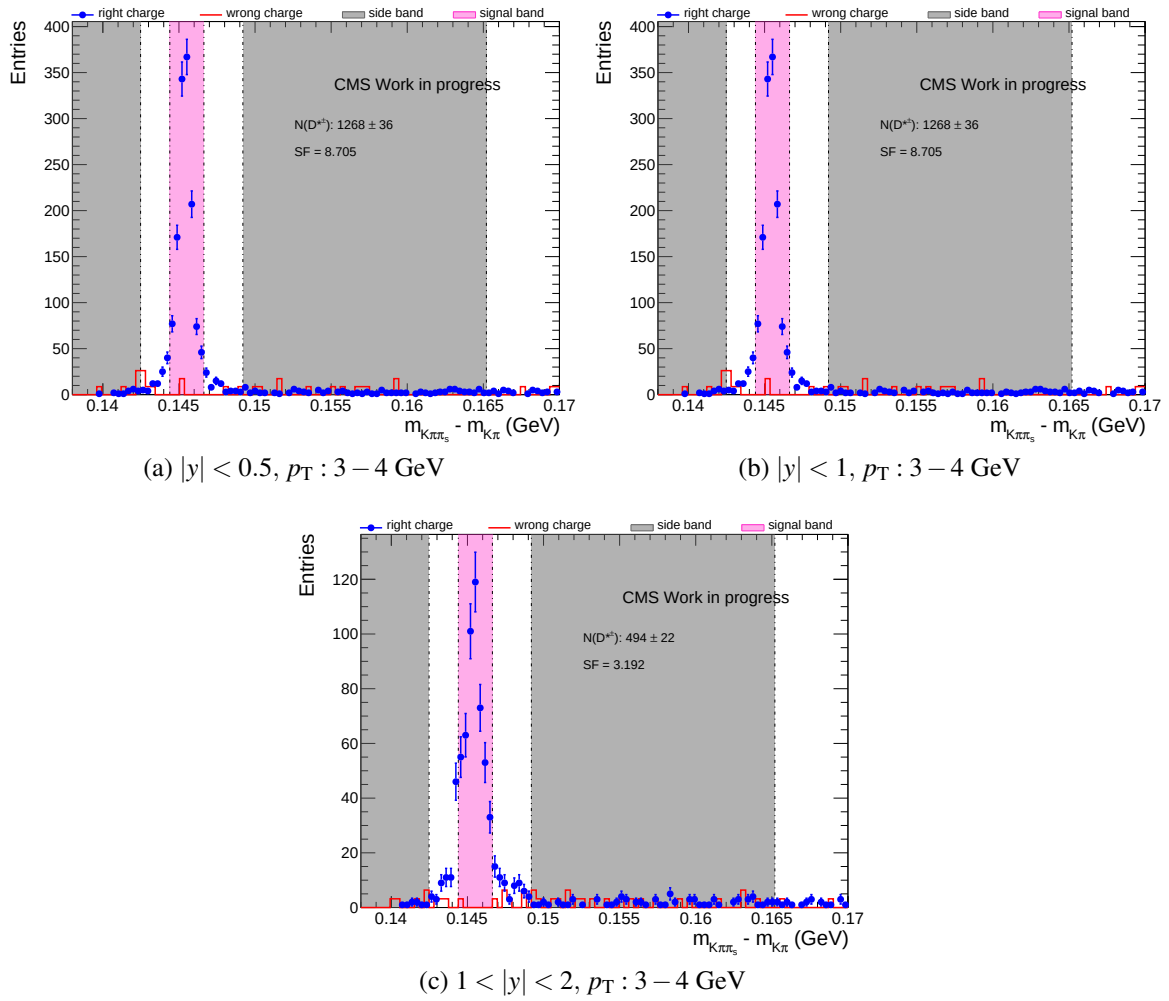


Fig. C.23 Signal extraction in MC for $p_T : 3 - 4$ GeV and $0 < |y| < 2$

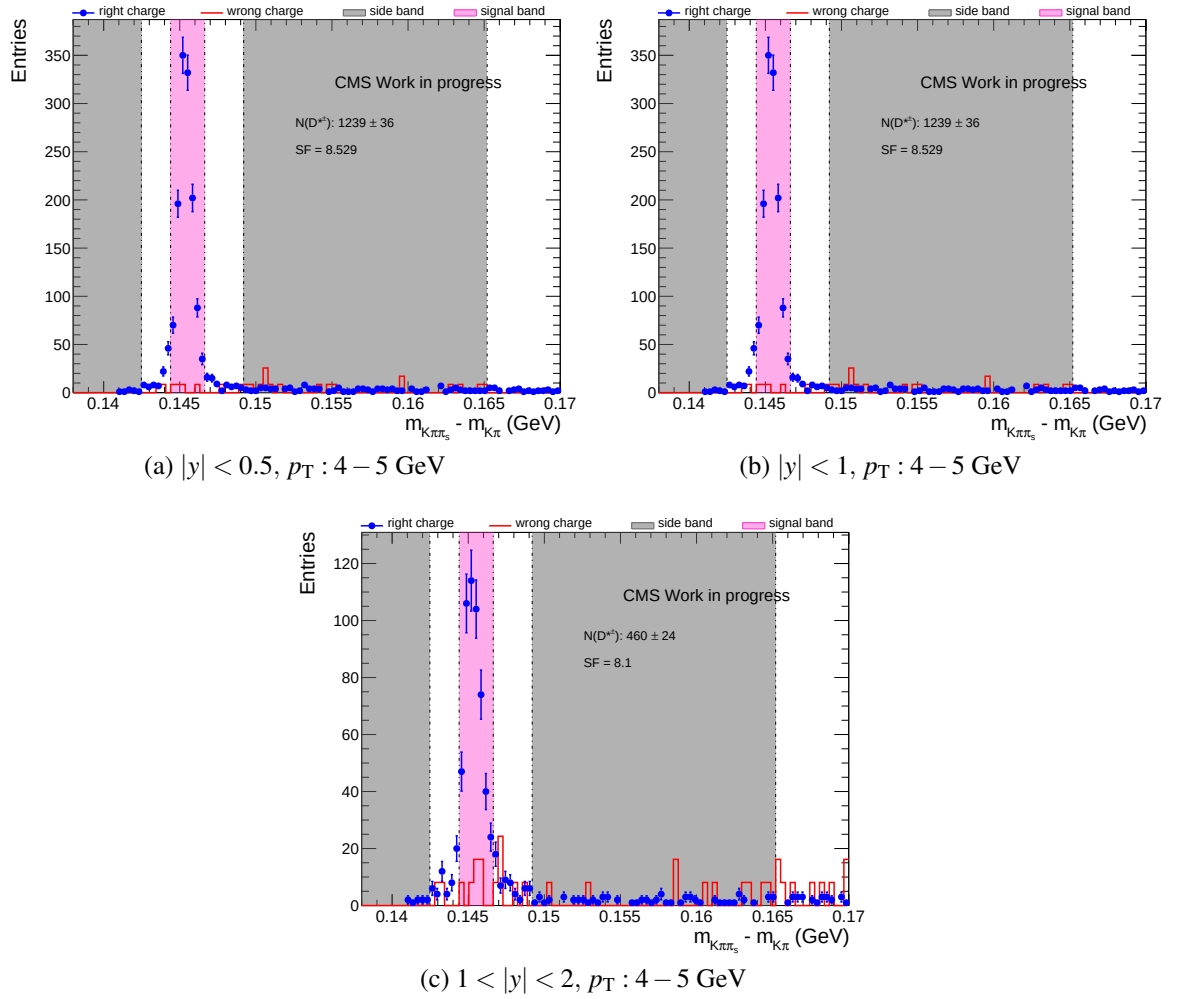


Fig. C.24 Signal extraction in MC for $p_T : 4 - 5$ GeV and $0 < |y| < 2$

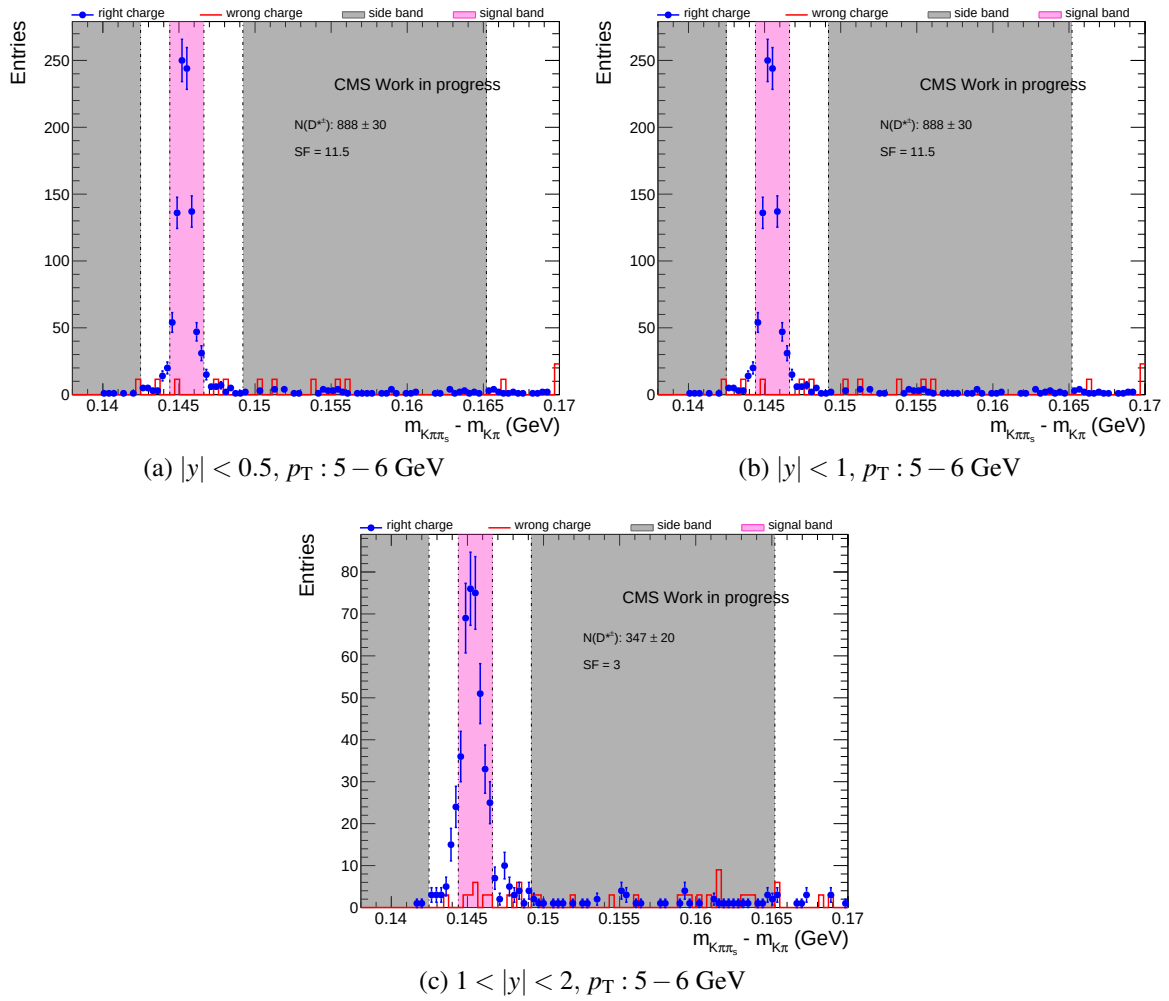


Fig. C.25 Signal extraction in MC for $p_T : 5 - 6$ GeV and $0 < |y| < 2$

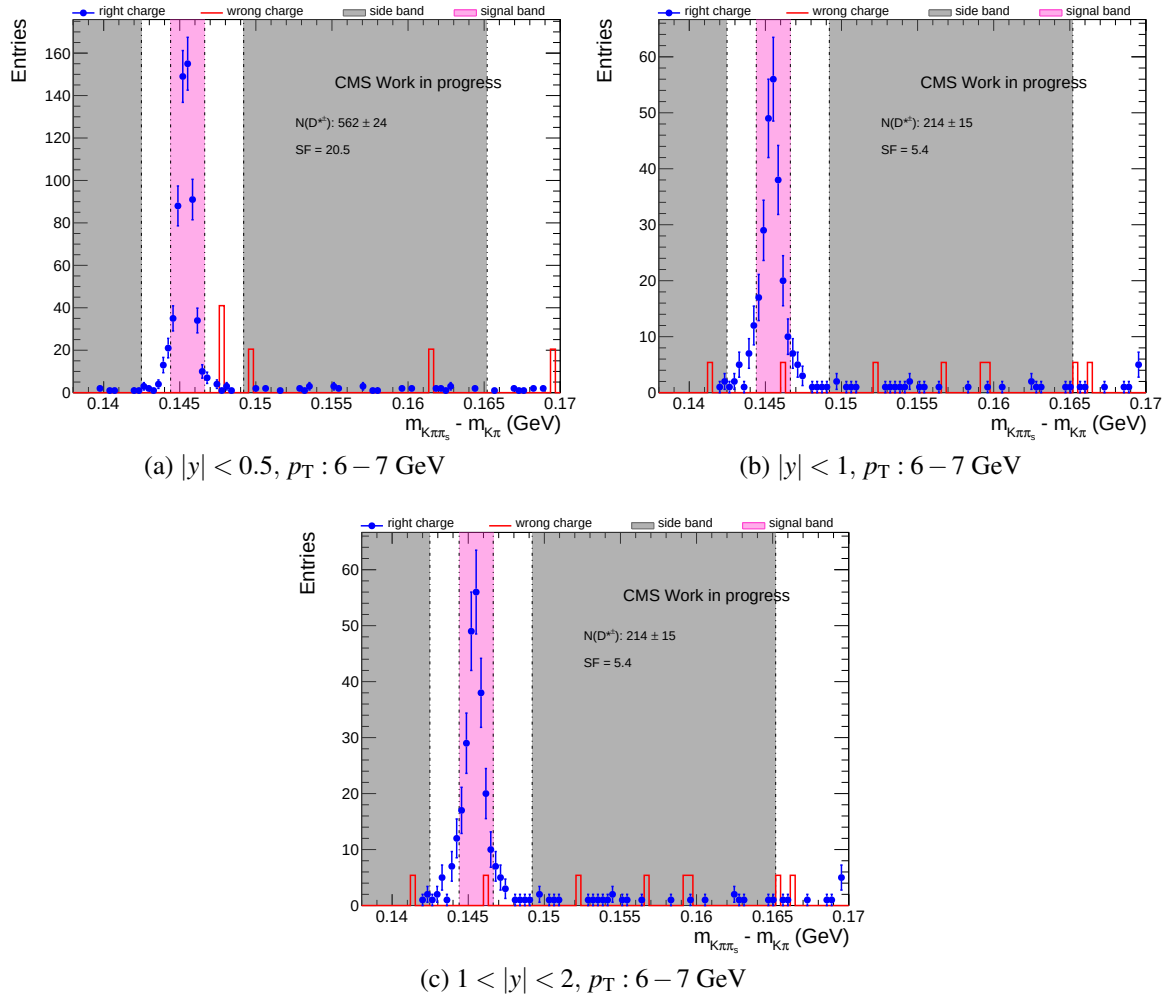


Fig. C.26 Signal extraction in MC for $p_T : 6-7$ GeV and $0 < |y| < 2$

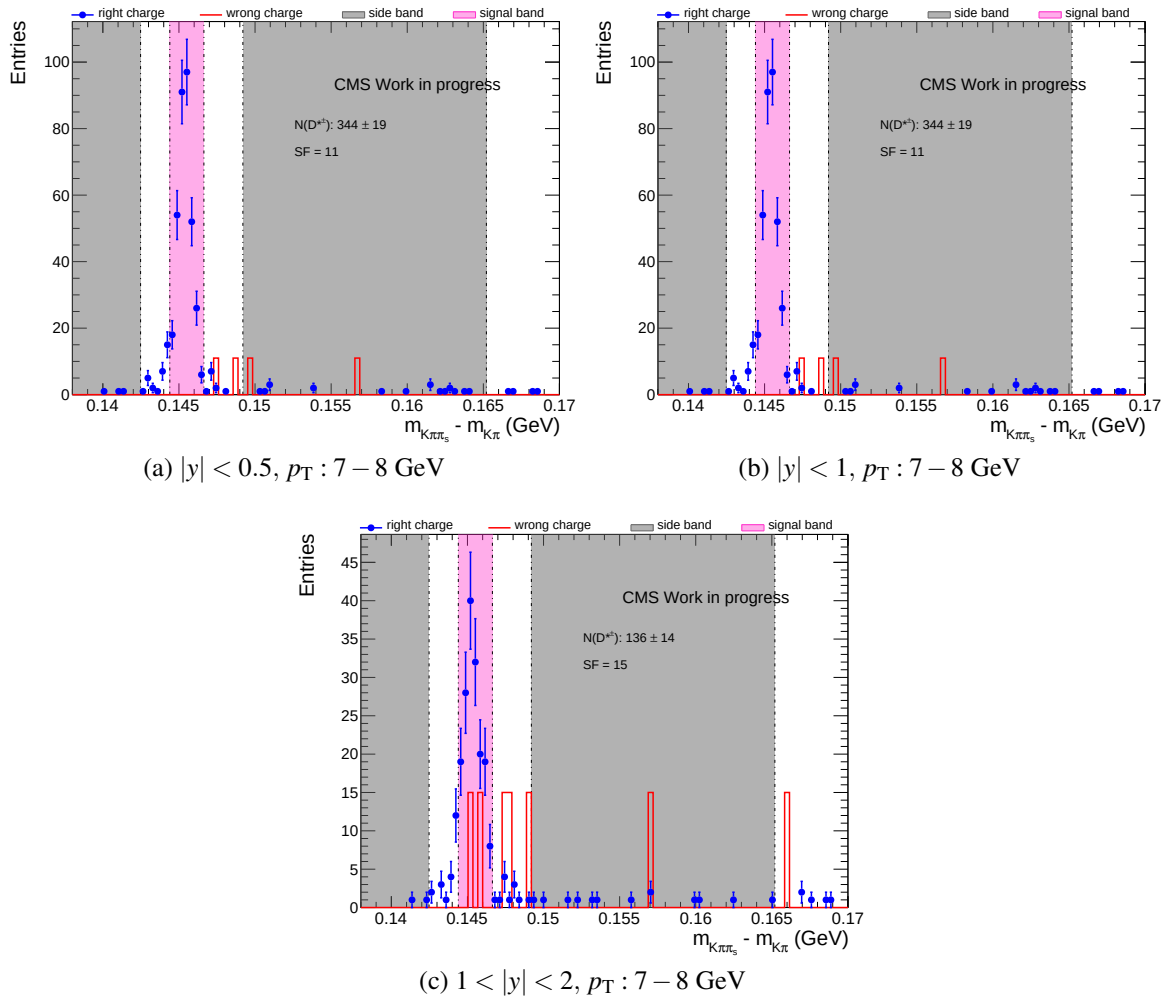


Fig. C.27 Signal extraction in MC for $p_T : 7 - 8$ GeV and $0 < |y| < 2$

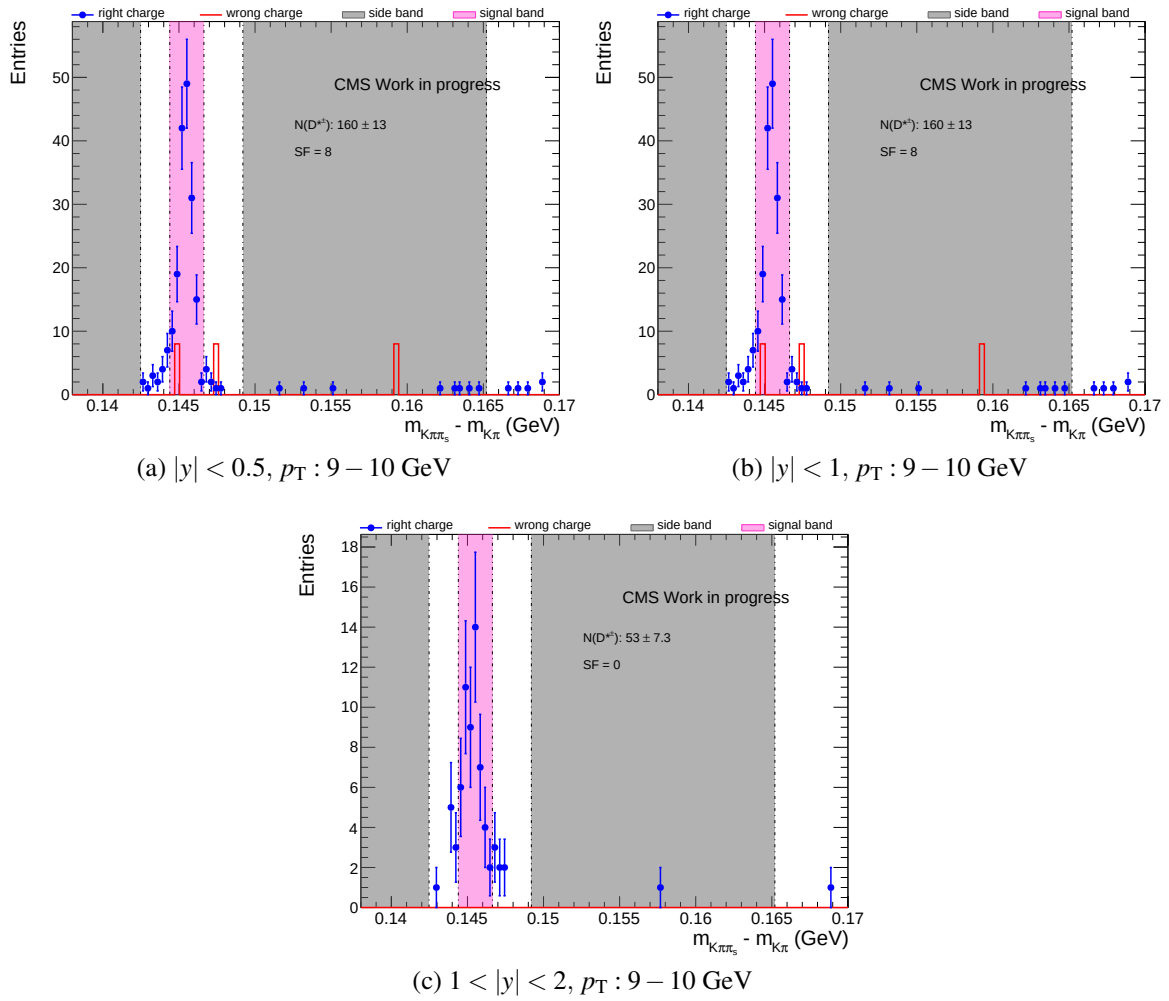


Fig. C.29 Signal extraction in MC for $p_T : 9 - 10$ GeV and $0 < |y| < 2$

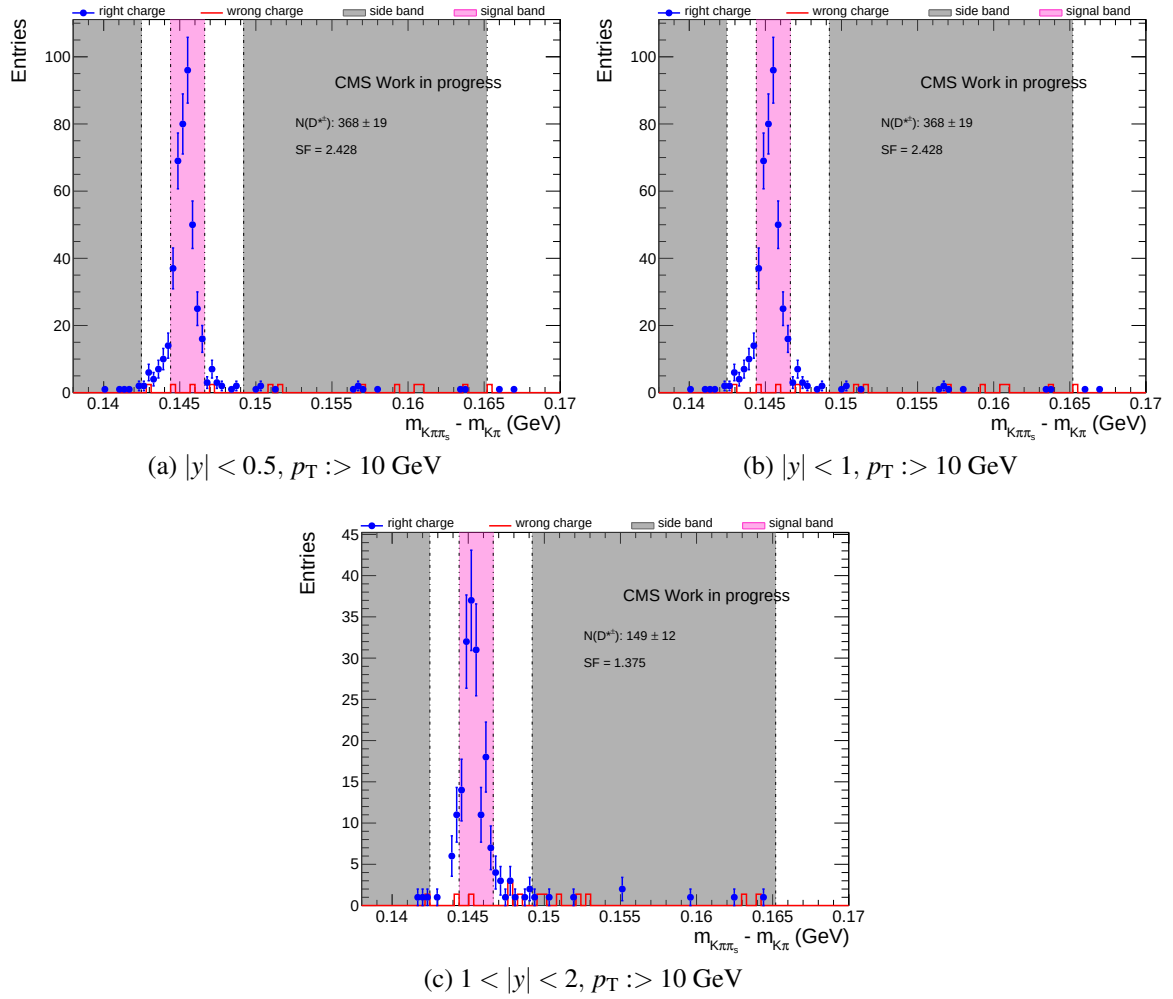


Fig. C.30 Signal extraction in MC for $p_T > 10$ GeV and $0 < |y| < 2$

C.3 Separation of Charm and Beauty in D^* Final States in Bins of p_T and Rapidity

In the following the charm beauty separation for the signal bins shown in Section C.1. First, based on dca and then followed by the method uses $dca \cdot dl_{sig}$.

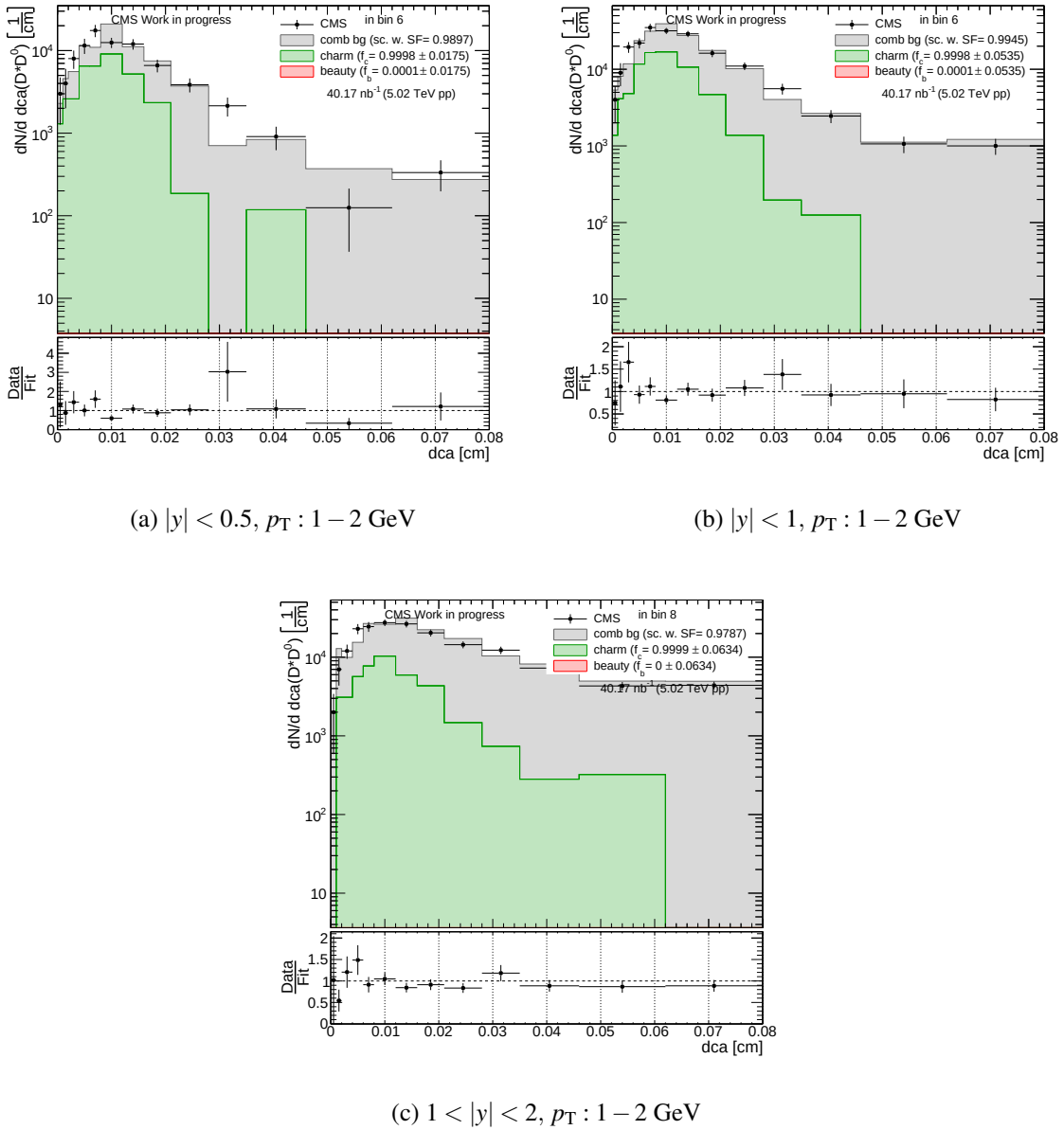


Fig. C.31 Charm beauty separation based on dca for $p_T : 1 - 2$ GeV and $0 < |y| < 2$

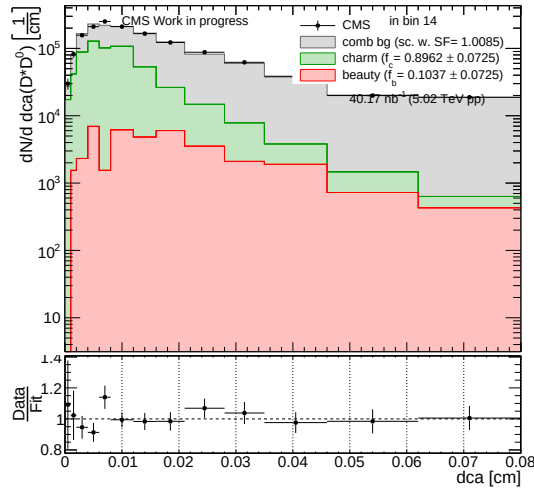
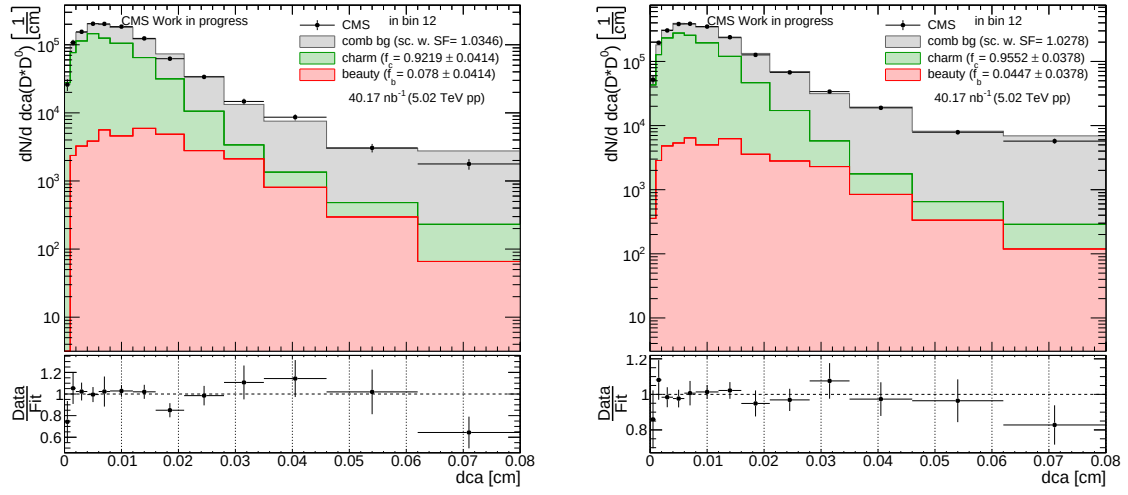
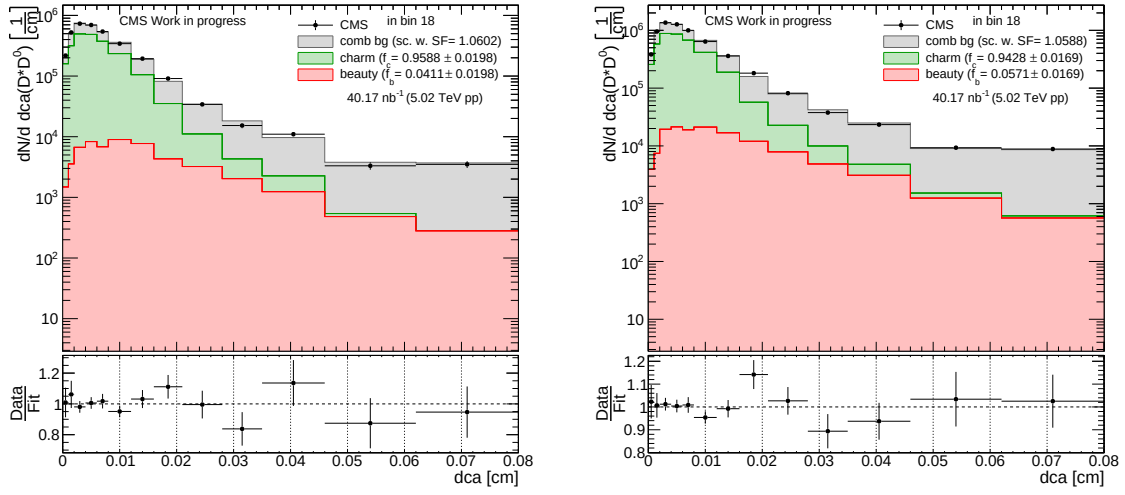
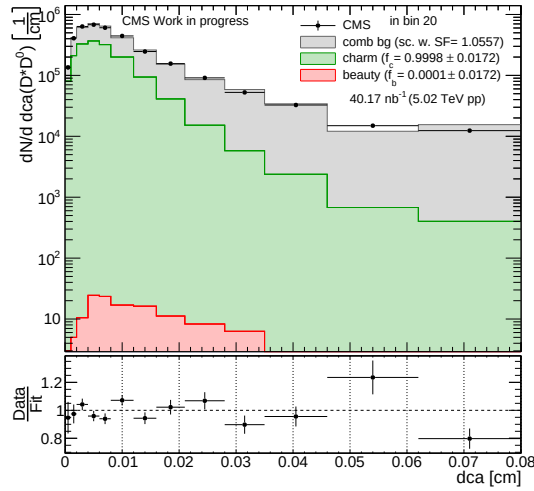

(c) $1 < |y| < 2, p_T : 2 - 3$ GeV

Fig. C.32 Charm beauty separation based on dca for $p_T : 2 - 3$ GeV and $0 < |y| < 2$



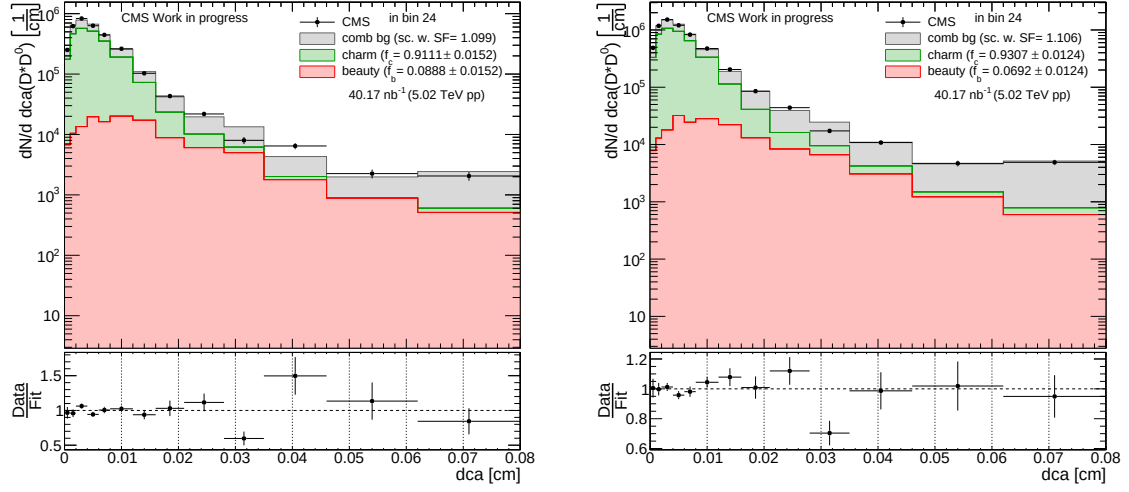
(a) $|y| < 0.5$, $p_T : 3 - 4$ GeV

(b) $|y| < 1$, $p_T : 3 - 4$ GeV



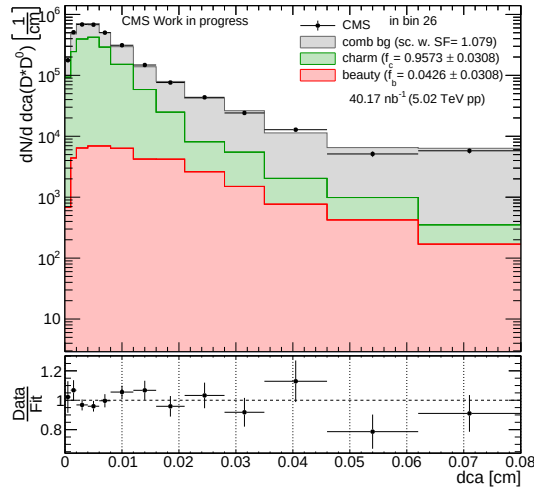
(c) $1 < |y| < 2$, $p_T : 3 - 4$ GeV

Fig. C.33 Charm beauty separation based on dca for $p_T : 3 - 4$ GeV and $0 < |y| < 2$



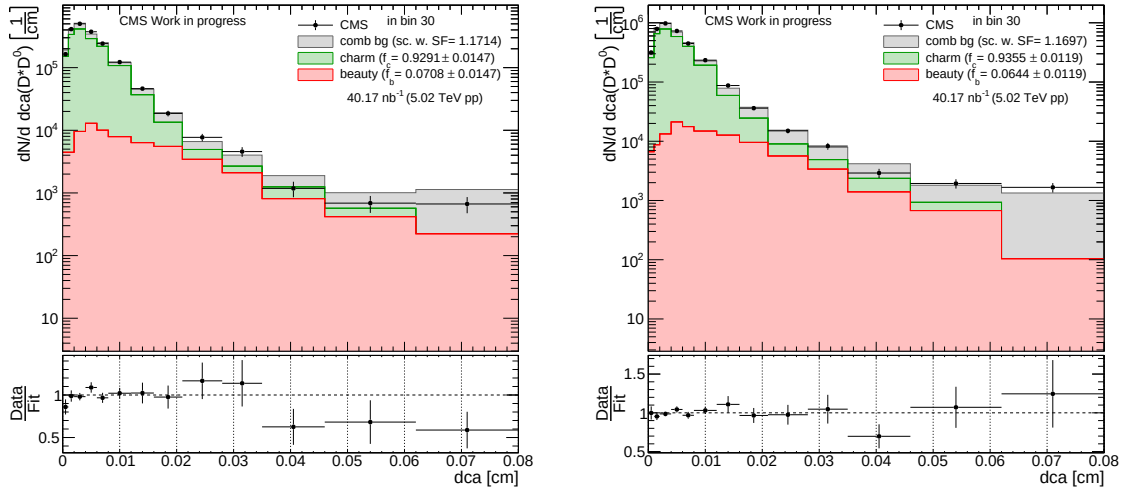
(a) $|y| < 0.5$, $p_T : 4-5$ GeV

(b) $|y| < 1$, $p_T : 4-5$ GeV



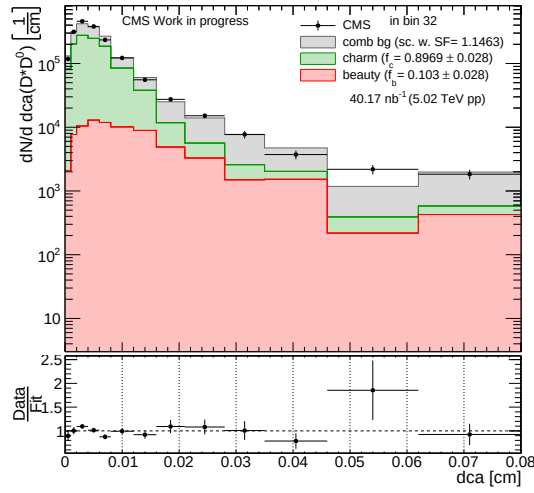
(c) $1 < |y| < 2$, $p_T : 4-5$ GeV

Fig. C.34 Charm beauty separation based on dca for $p_T : 4-5$ GeV and $0 < |y| < 2$



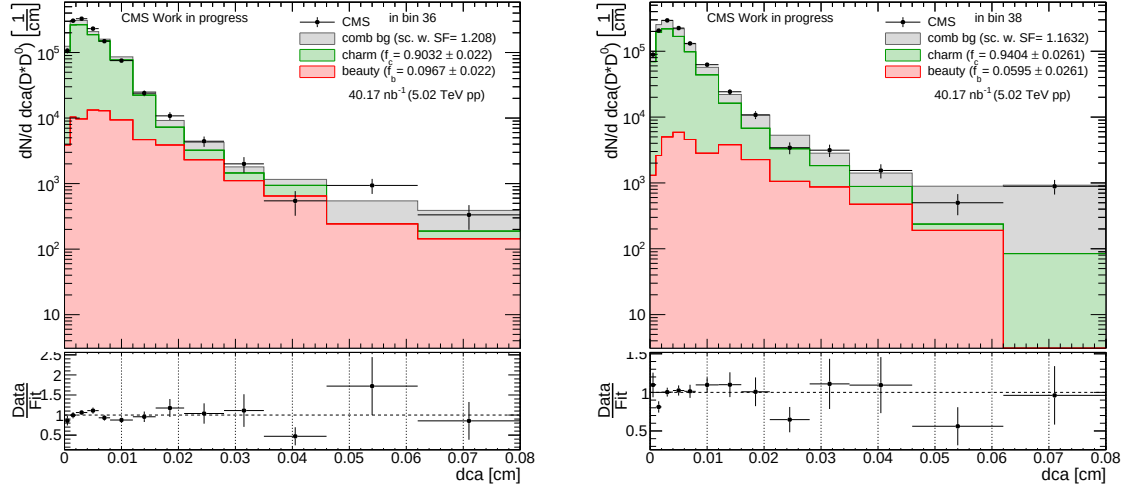
(a) $|y| < 0.5$, $p_T : 5 - 6 \text{ GeV}$

(b) $|y| < 1$, $p_T : 5 - 6 \text{ GeV}$



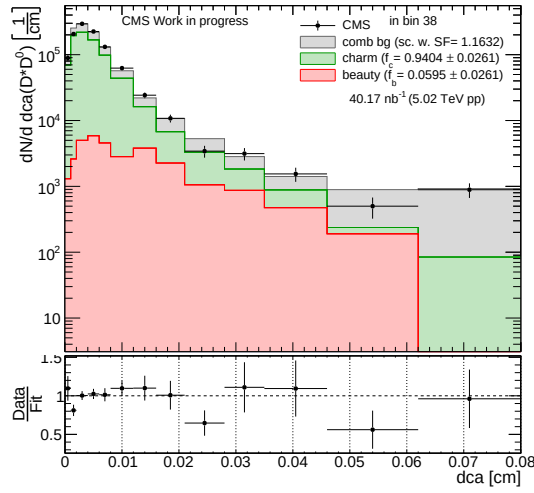
(c) $1 < |y| < 2$, $p_T : 5 - 6 \text{ GeV}$

Fig. C.35 Charm beauty separation based on dca for $p_T : 5 - 6 \text{ GeV}$ and $0 < |y| < 2$



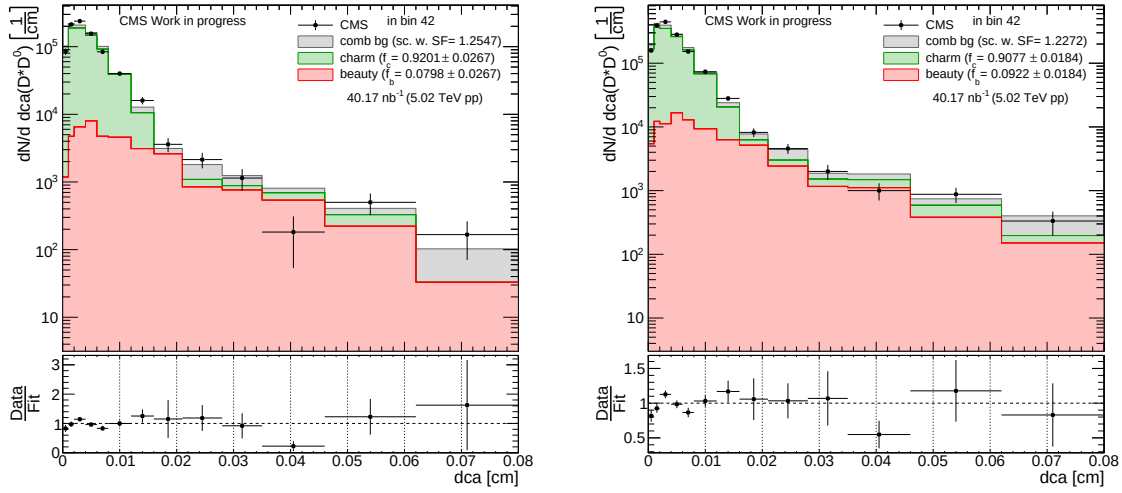
(a) $|y| < 0.5$, $p_T : 6 - 7$ GeV

(b) $|y| < 1$, $p_T : 6 - 7$ GeV



(c) $1 < |y| < 2$, $p_T : 6 - 7$ GeV

Fig. C.36 Charm beauty separation based on dca for $p_T : 6 - 7$ GeV and $0 < |y| < 2$


(a) $|y| < 0.5$, $p_T : 7 - 8$ GeV

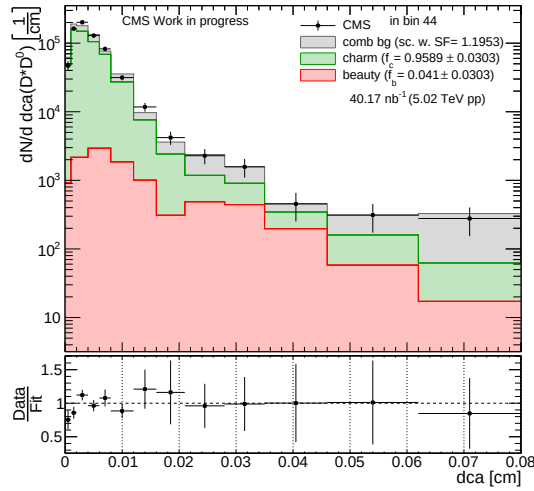
(b) $|y| < 1$, $p_T : 7 - 8$ GeV

(c) $1 < |y| < 2$, $p_T : 7 - 8$ GeV

Fig. C.37 Charm beauty separation based on dca for $p_T : 7 - 8$ GeV and $0 < |y| < 2$

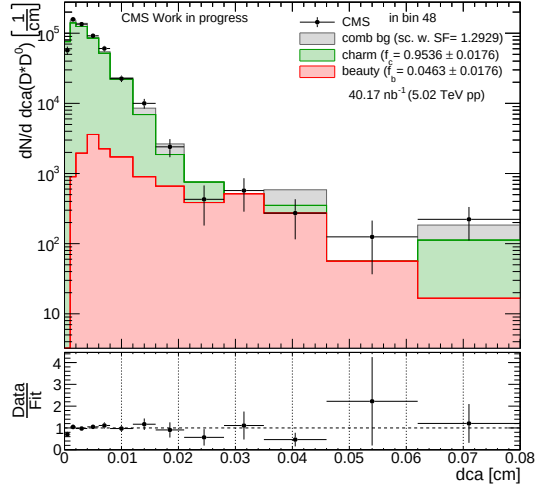
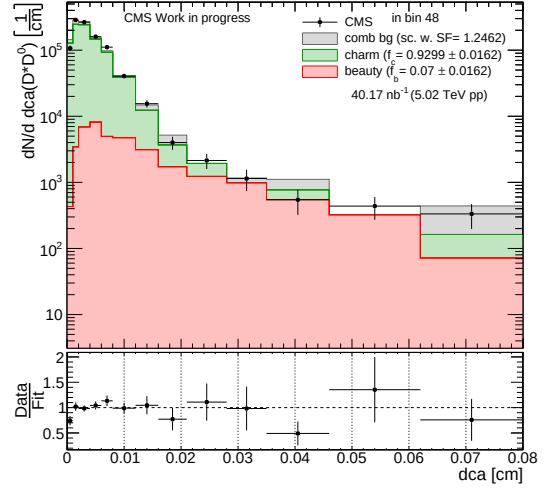
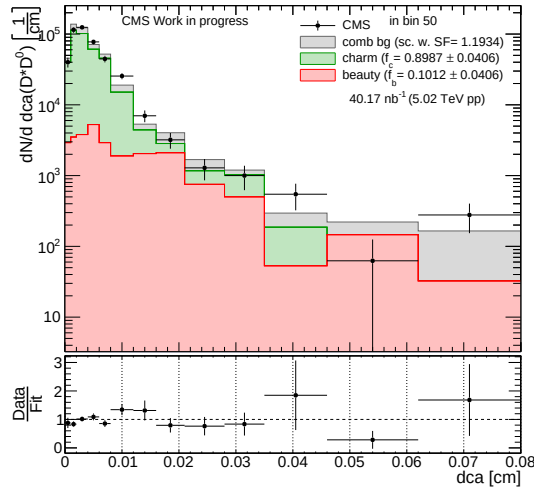
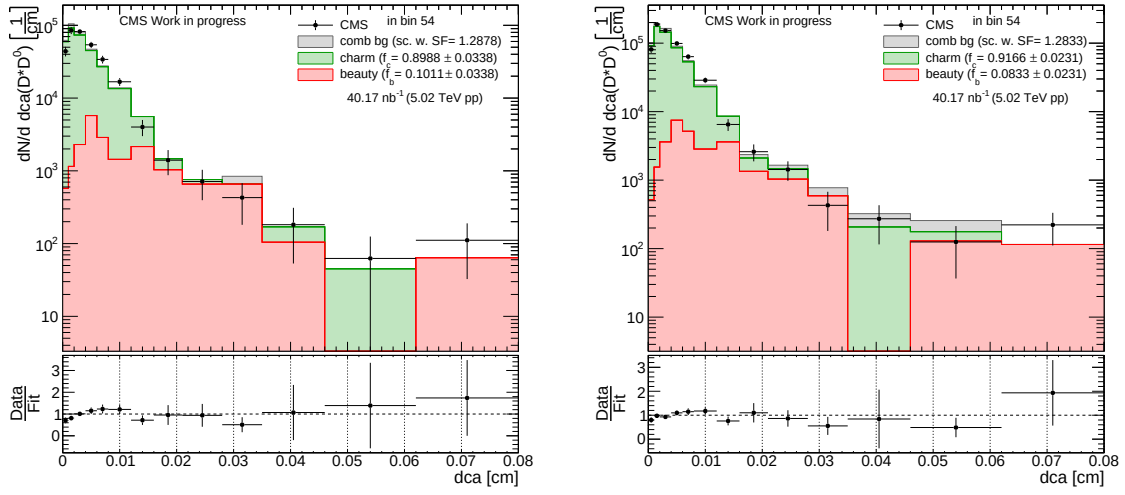

(a) $|y| < 0.5$, $p_T : 8 - 9$ GeV

(b) $|y| < 1$, $p_T : 8 - 9$ GeV

(c) $1 < |y| < 2$, $p_T : 8 - 9$ GeV

Fig. C.38 Charm beauty separation based on dca for $p_T : 8 - 9$ GeV and $0 < |y| < 2$


(a) $|y| < 0.5$, $p_T : 9 - 10$ GeV

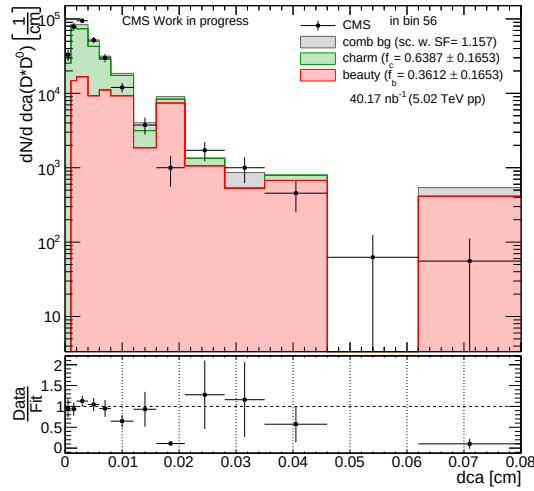
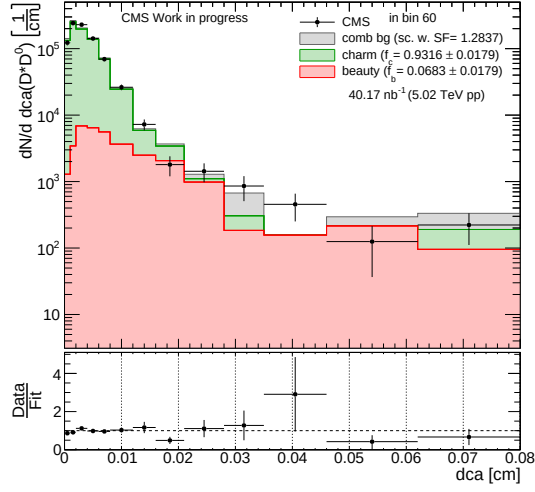
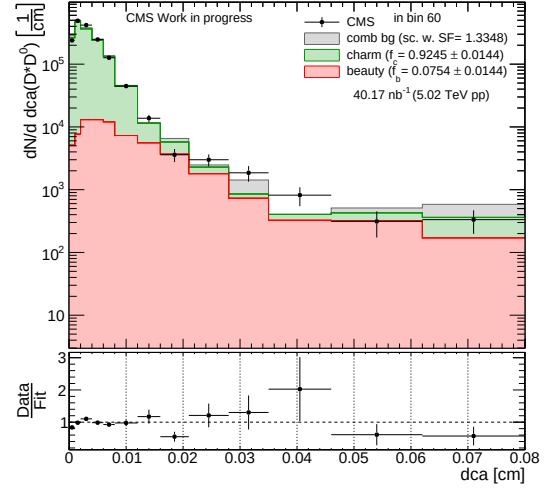
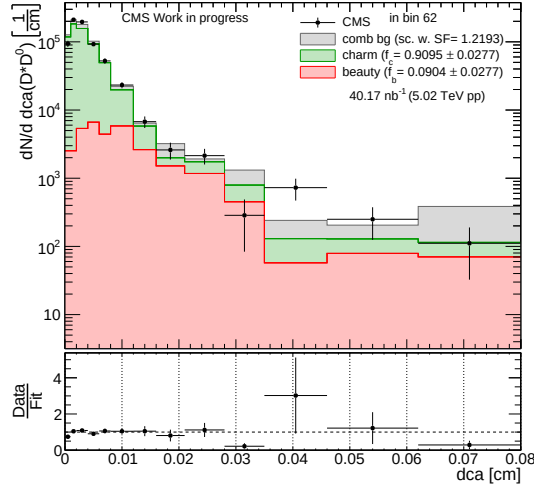
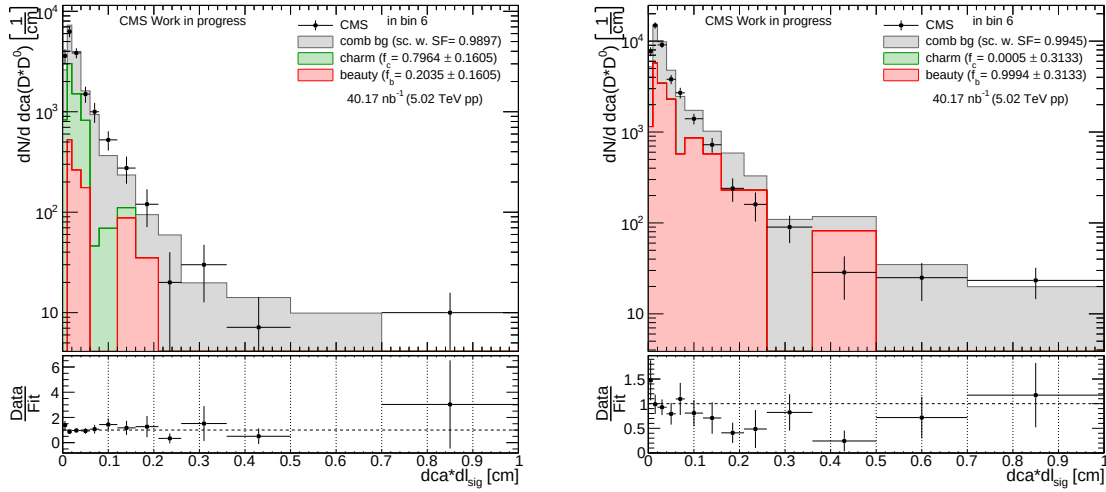
(b) $|y| < 1$, $p_T : 9 - 10$ GeV

(c) $1 < |y| < 2$, $p_T : 9 - 10$ GeV

Fig. C.39 Charm beauty separation based on dca for $p_T : 9 - 10$ GeV and $0 < |y| < 2$

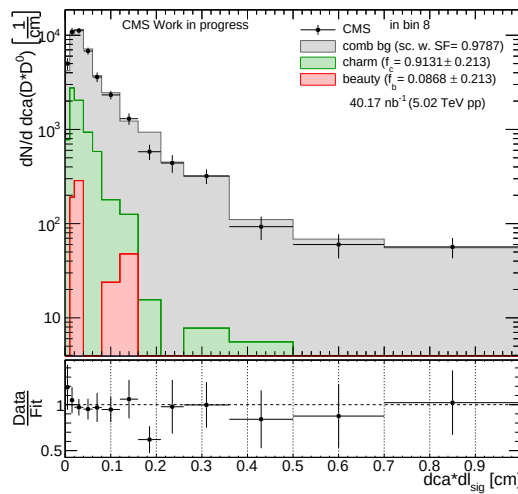
(a) $|y| < 0.5$, $p_T > 10$ GeV(b) $|y| < 1$, $p_T > 10$ GeV(c) $1 < |y| < 2$, $p_T > 10$ GeVFig. C.40 Charm beauty separation based on dca for $p_T > 10$ GeV and $0 < |y| < 2$

Now in the following the charm beauty separation using the parameter $dca \cdot dl_{sig}$ to distinguish between prompt and non-prompt D^* .



(a) $|y| < 0.5$, $p_T : 1 - 2 \text{ GeV}$

(b) $|y| < 1$, $p_T : 1 - 2 \text{ GeV}$



(c) $1 < |y| < 2$, $p_T : 1 - 2 \text{ GeV}$

Fig. C.41 Charm beauty separation based on $dca \cdot dl_{sig}$ for $p_T : 1 - 2 \text{ GeV}$ and $0 < |y| < 2$

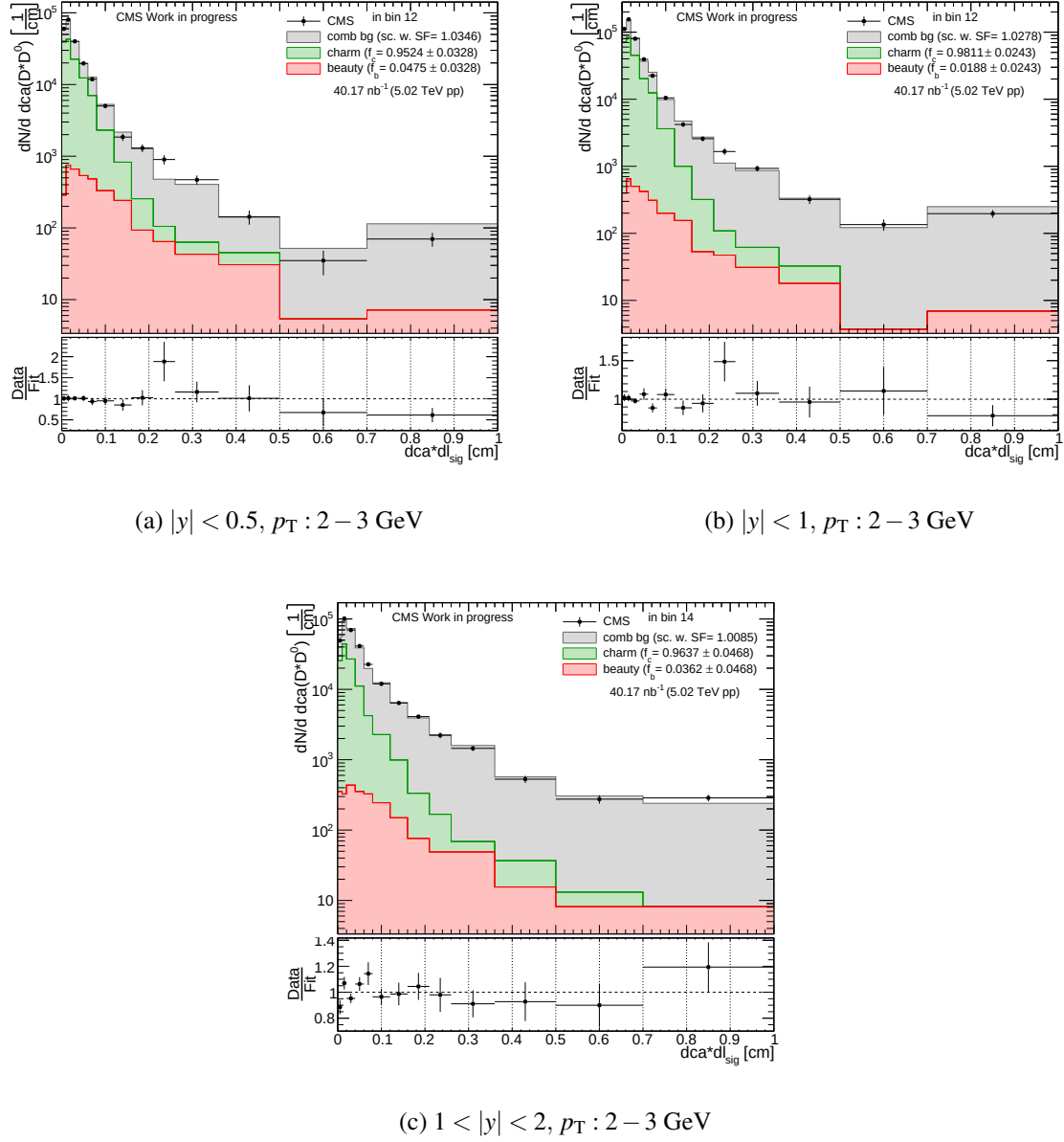


Fig. C.42 Charm beauty separation based on $dca \cdot dl_{sig}$ for $p_T : 2 - 3$ GeV and $0 < |y| < 2$

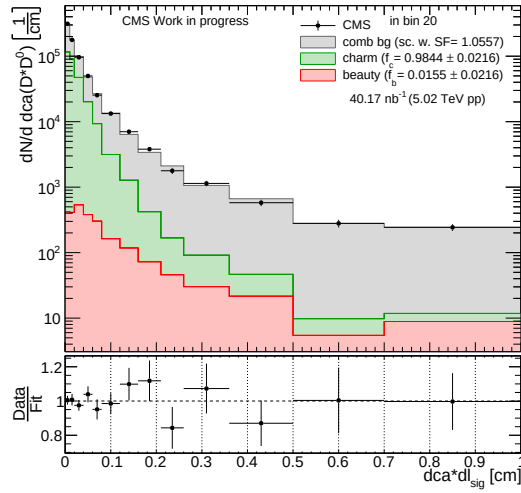
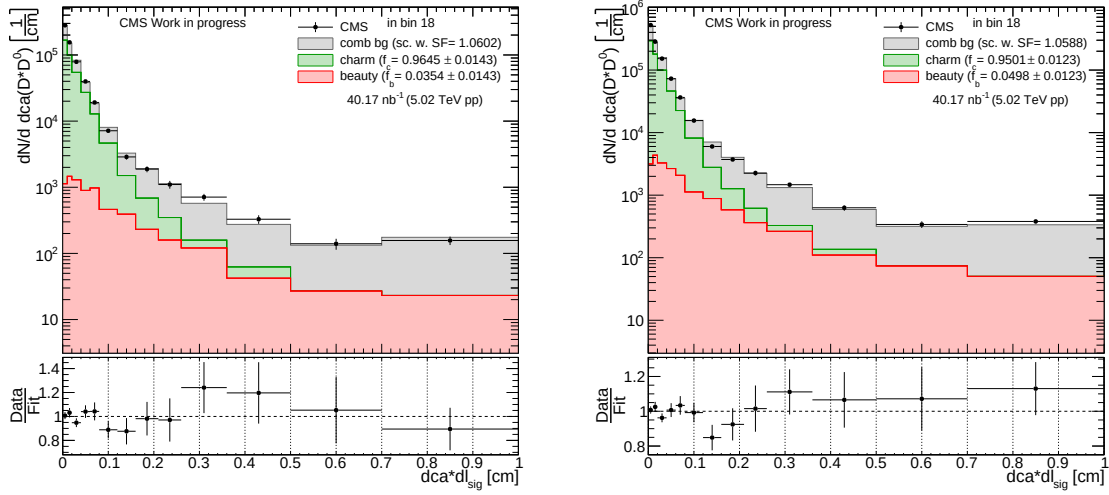
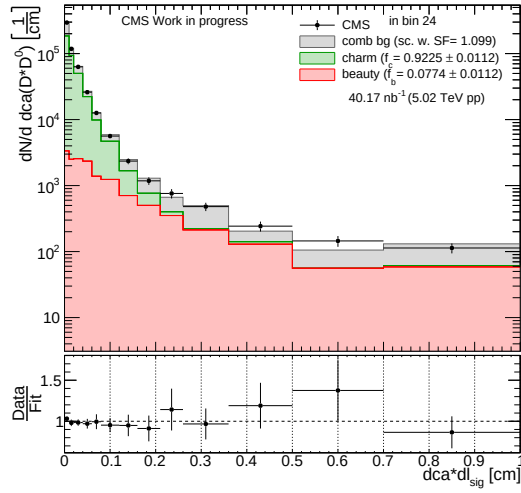
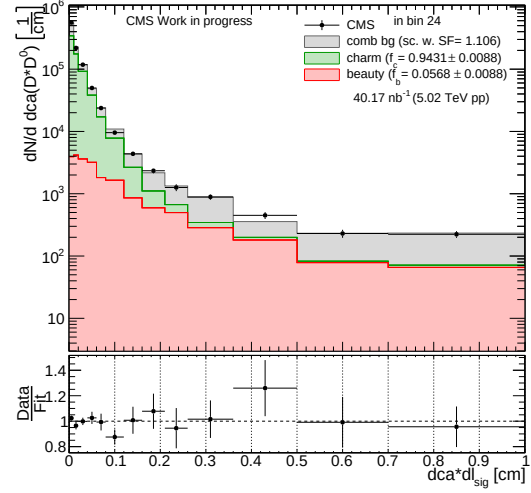


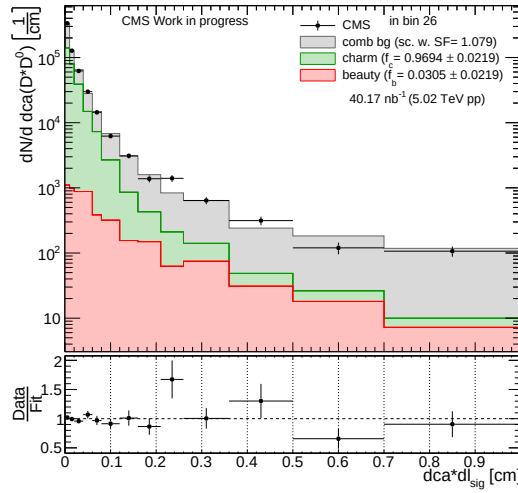
Fig. C.43 Charm beauty separation based on $dca \cdot dl_{sig}$ for $p_T : 3 - 4$ GeV and $0 < |y| < 2$



(a) $|y| < 0.5$, $p_T : 4 - 5$ GeV



(b) $|y| < 1$, $p_T : 4 - 5$ GeV



(c) $1 < |y| < 2$, $p_T : 4 - 5$ GeV

Fig. C.44 Charm beauty separation based on $dca \cdot dl_{sig}$ for $p_T : 4 - 5$ GeV and $0 < |y| < 2$

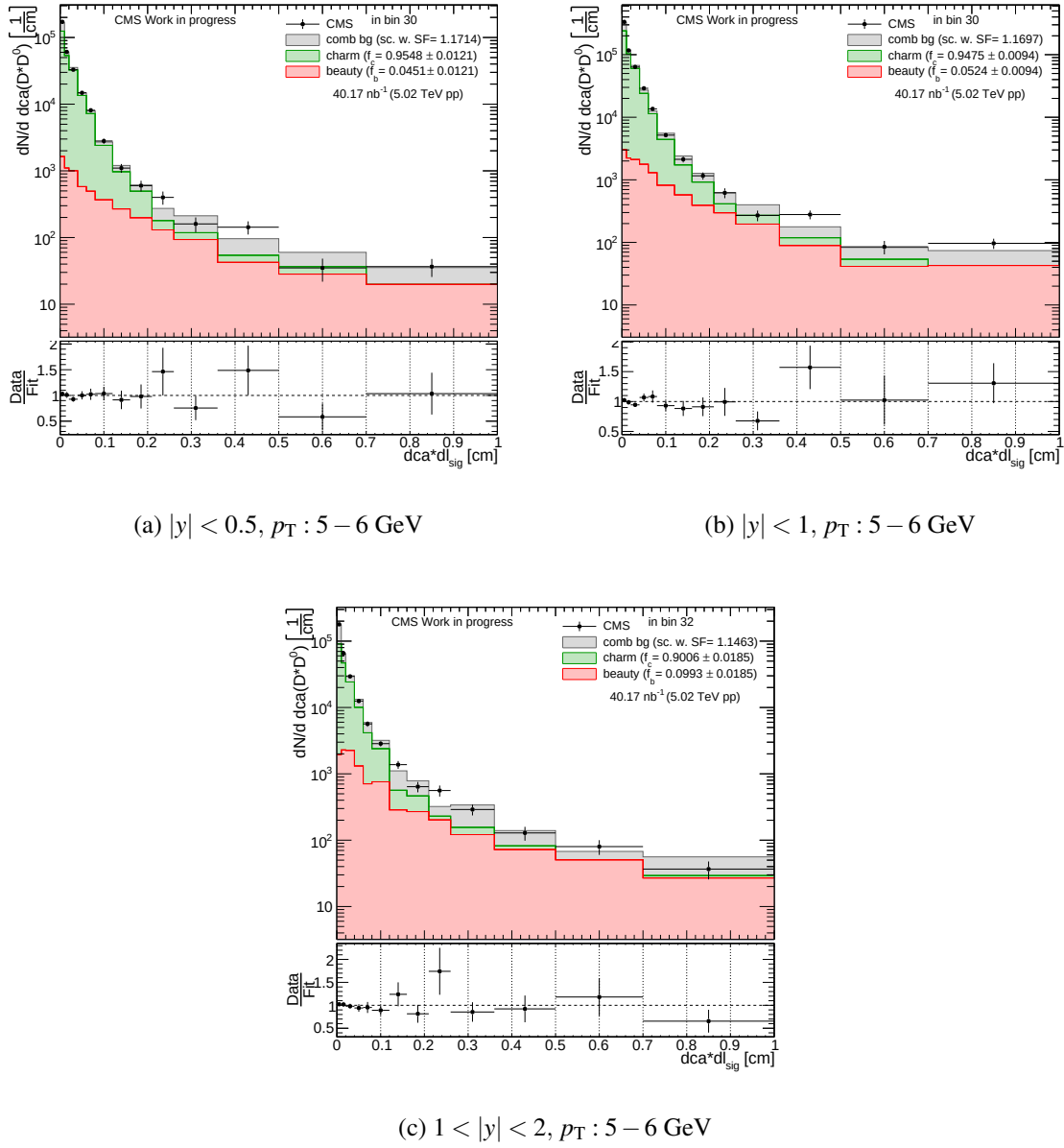
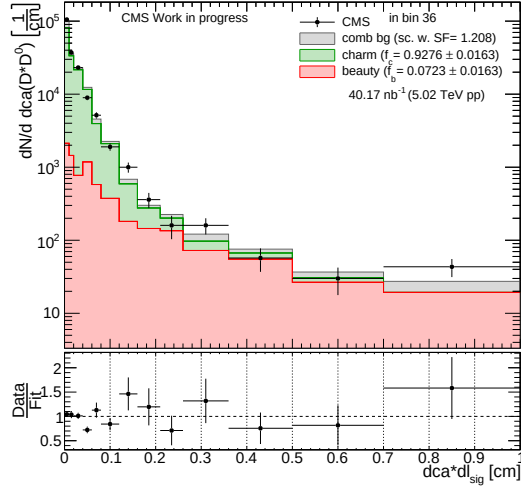
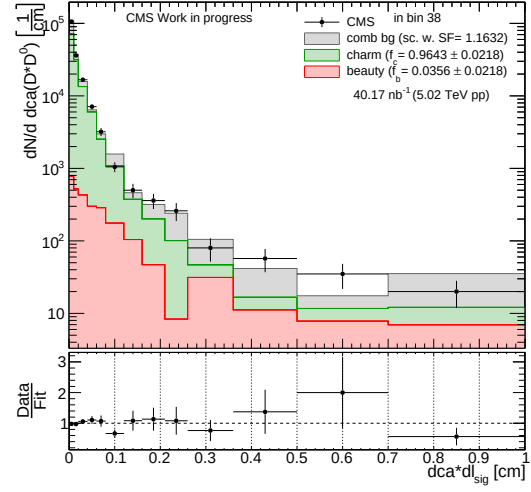


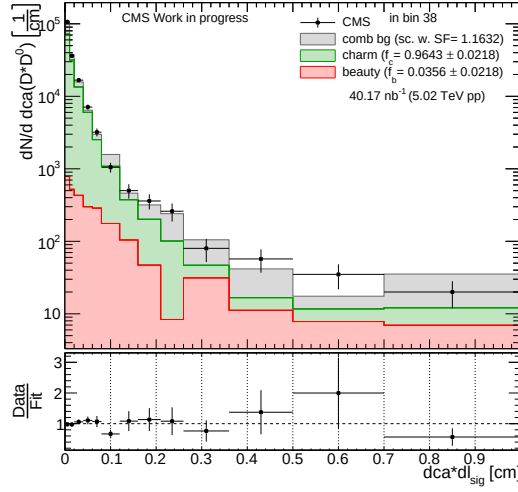
Fig. C.45 Charm beauty separation based on $dca \cdot dl_{sig}$ for $p_T : 5 - 6 \text{ GeV}$ and $0 < |y| < 2$



(a) $|y| < 0.5$, $p_T : 6-7$ GeV

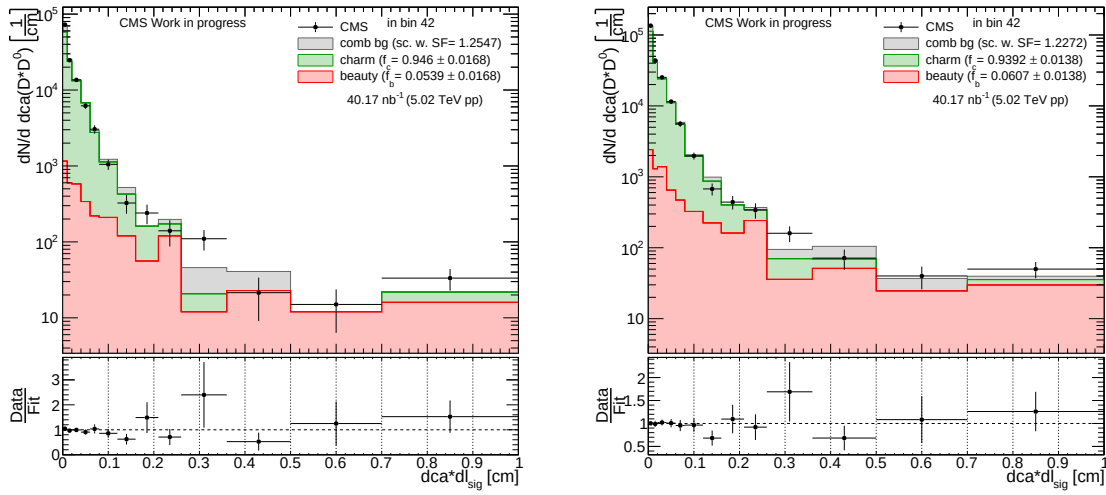


(b) $|y| < 1$, $p_T : 6-7$ GeV



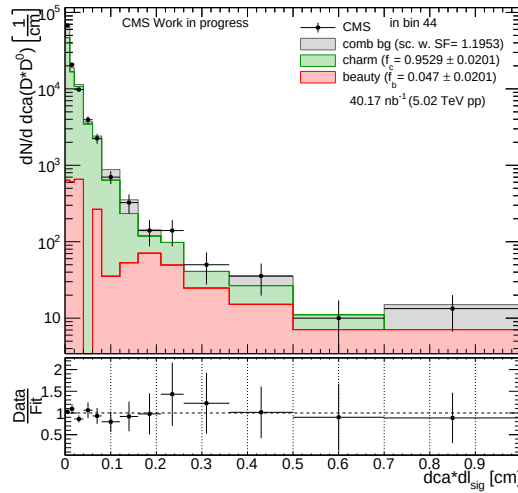
(c) $1 < |y| < 2$, $p_T : 6-7$ GeV

Fig. C.46 Charm beauty separation based on $dca \cdot dl_{sig}$ for $p_T : 6-7$ GeV and $0 < |y| < 2$



(a) $|y| < 0.5$, $p_T : 7 - 8$ GeV

(b) $|y| < 1$, $p_T : 7 - 8$ GeV



(c) $1 < |y| < 2$, $p_T : 7 - 8$ GeV

Fig. C.47 Charm beauty separation based on $dca \cdot dl_{sig}$ for $p_T : 7 - 8$ GeV and $0 < |y| < 2$

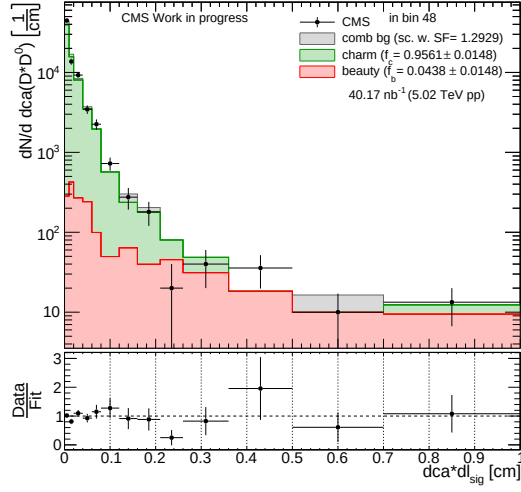
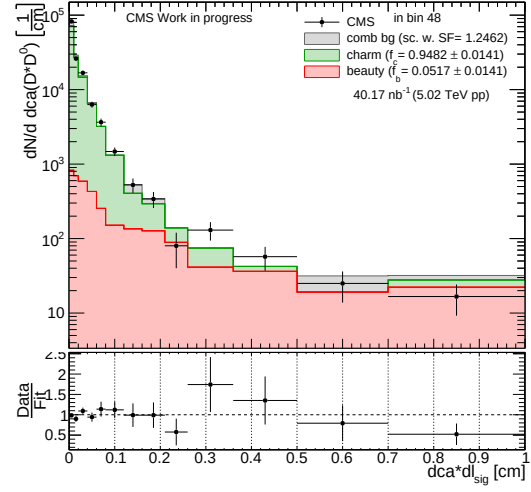
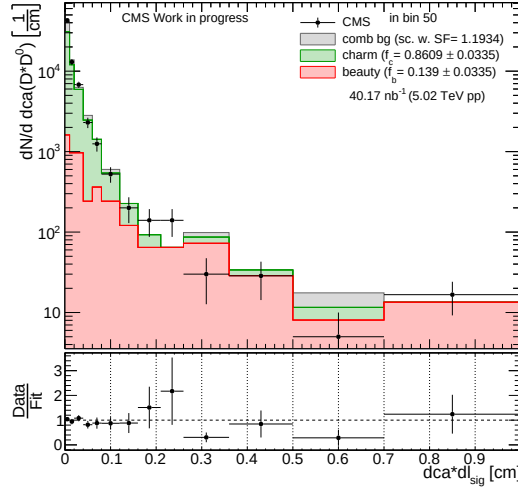

(a) $|y| < 0.5$, $p_T : 8 - 9$ GeV

(b) $|y| < 1$, $p_T : 8 - 9$ GeV

(c) $1 < |y| < 2$, $p_T : 8 - 9$ GeV

Fig. C.48 Charm beauty separation based on $dca \cdot dl_{sig}$ for $p_T : 8 - 9$ GeV and $0 < |y| < 2$

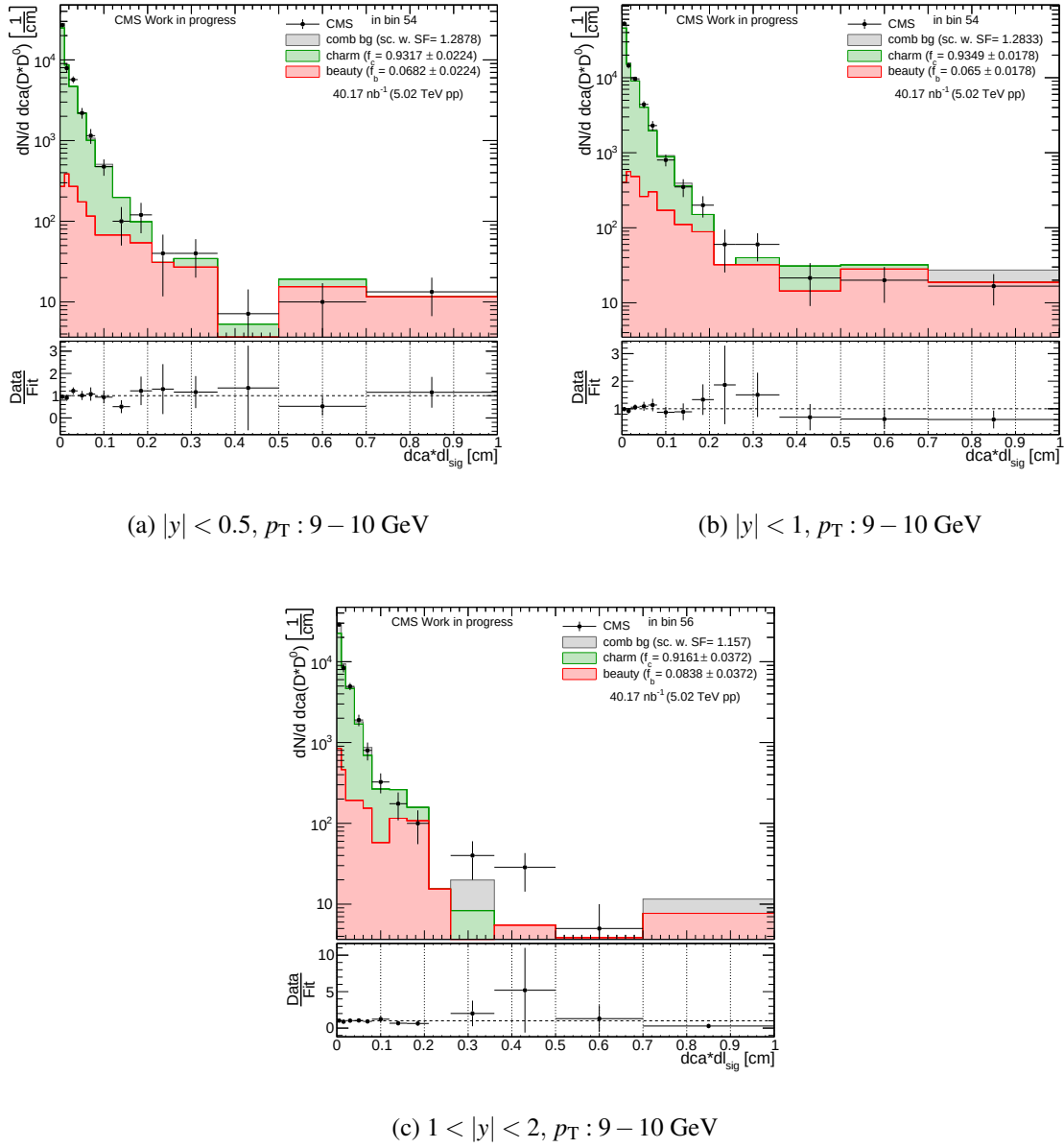
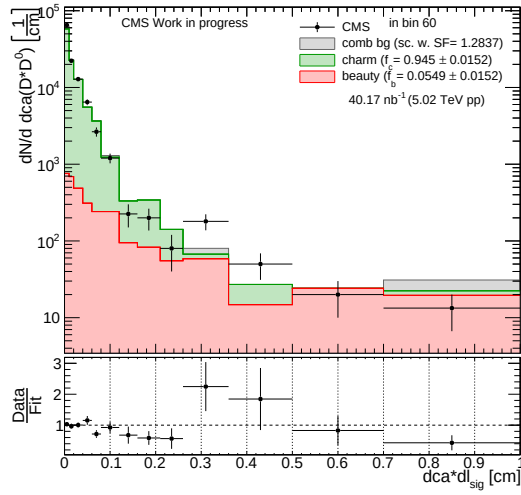
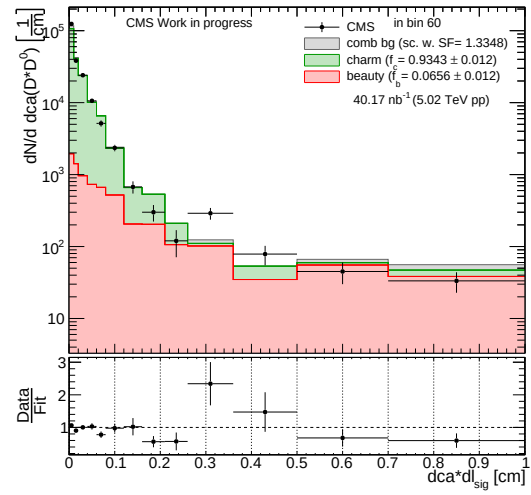


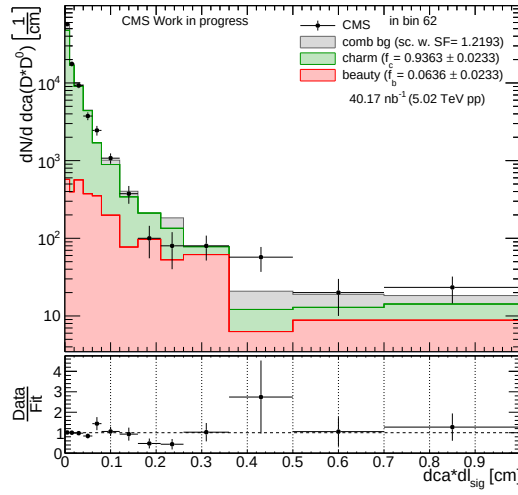
Fig. C.49 Charm beauty separation based on $dca \cdot dl_{sig}$ for $p_T : 9 - 10$ GeV and $0 < |y| < 2$



(a) $|y| < 0.5$, $p_T > 10$ GeV



(b) $|y| < 1$, $p_T > 10$ GeV



(c) $1 < |y| < 2$, $p_T > 10$ GeV

Fig. C.50 Charm beauty separation based on $\text{dca} \cdot dl_{\text{sig}}$ for $p_T > 10$ GeV and $0 < |y| < 2$

C.4 Investigate of Correlations - A Different Ansatz for the Charm Beauty Separation

In the following a selection of correlation plots used to investigate a different method for the charm beauty separation

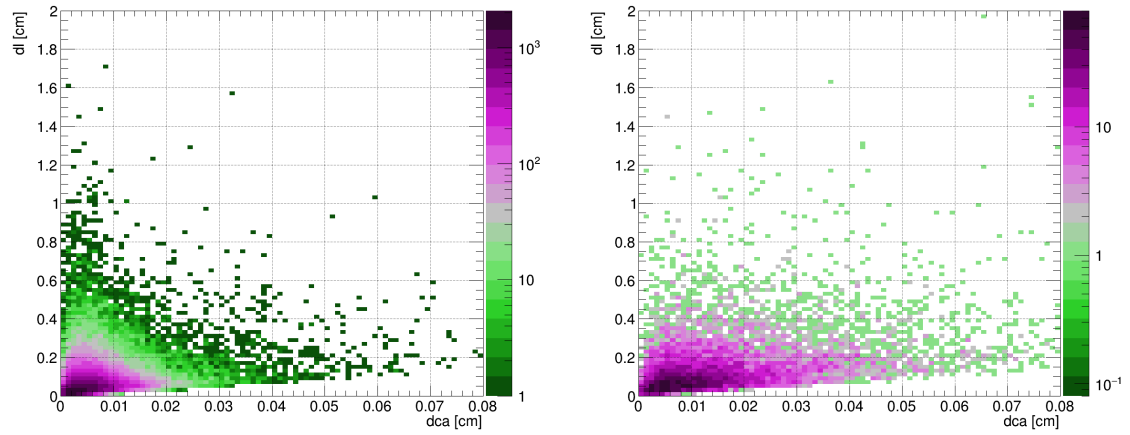


Fig. C.51 Correlation plots of the decay length and the dca from the D^0 for charm (left) and beauty (right) for $p_T > 3.5$ GeV and $|y| < 2.5$.

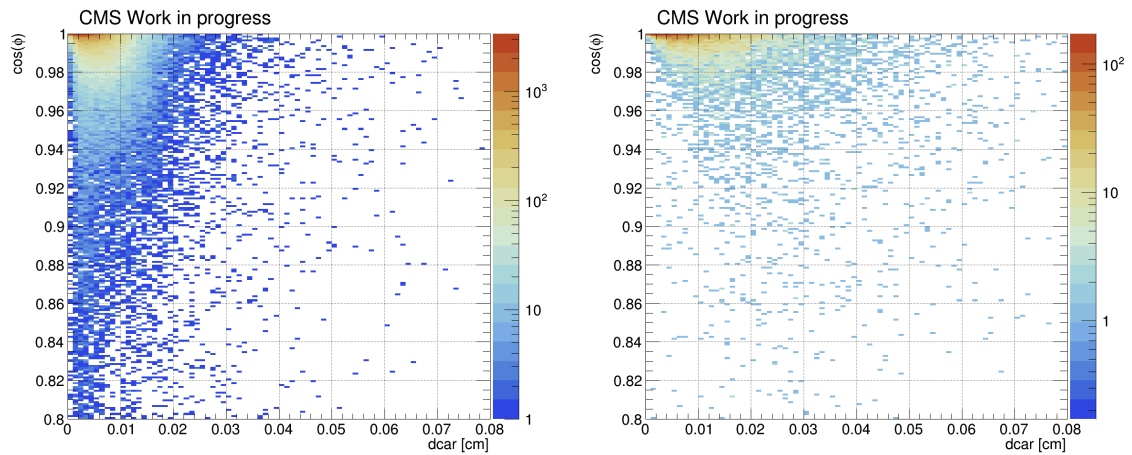


Fig. C.52 Correlation plots of the decay angle ϕ and the dca from the D^0 for charm (left) and beauty (right) for $p_T > 3.5$ GeV and $|y| < 2.5$.

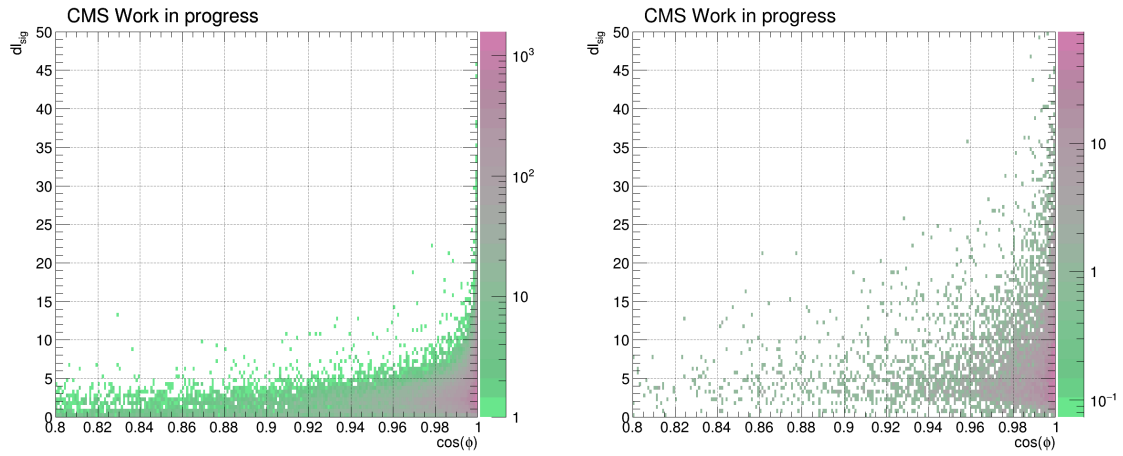


Fig. C.53 Correlation plots of the decay length significance and the the angle ϕ from the D^0 for charm (left) and beauty (right) for $p_T > 3.5$ GeV and $|y| < 2.5$.

List of figures

1.1	Top pair production as a function of center of mass energy	3
2.1	A summary of the elementary particles and their anti-particles of the Standard Model of particle physics.	6
2.2	SM Feynman diagrams representing interactions between fermions and bosons.	9
2.3	Gluon self-interaction	14
2.4	Observation of the gluon at the TASSO detector and the corresponding Feynman diagram for this process.	15
2.5	Factorization of the cross section calculation.	20
2.6	Drell-Yan process	21
2.7	NNPDF3.1 Parton distribution functions for different parton flavors.	22
2.8	Variation of the energy scales μ_R and μ_F	24
3.1	Schematic view of the LHC and its pre-accelerators at CERN.	34
3.2	Schematic view of the CMS and its sub-detectors.	35
3.3	Schematic view of the muon system.	37
3.4	Schematic view of the HCAL.	38
3.5	Schematic view of the ECAL.	39
3.6	Schematic view of the tracker system.	40
3.7	Schematic view of the pixel detector system.	41
3.8	The challenge of the alignment in pictures.	42
	(a) The challenge of the alignment ideal versus non-ideal.	42
	(b) The challenge of the alignment at different scenarios.	42
3.9	The challenge of the alignment.	42
3.10	The picture before and after the track based alignment procedure.	43
3.11	The weak mode twist.	44
3.12	CMS simulation of the weak mode <i>Twist</i>	44

3.13	Comparison of the impact of two different alignments (prompt reco vs UL) to a D meson analysis.	45
3.14	Comparison of the use of two different alignments (prompt reco vs UL) for a J/Ψ analysis.	46
4.1	An event display of proton-proton collisions in the center of the CMS detector and for that same event the view from the outer part of the detector.	51
4.2	Illustration of the particle flow reconstruction.	52
4.3	Illustration of a heavy-flavor jet reconstruction.	55
4.4	Illustration of the architecture of the L1 trigger system.	56
4.5	Data flow between the different Tier centers.	57
4.6	A typical event of a hadron-hadron collision simulated by a Monte-Carlo event generator.	59
5.1	Feynman diagrams for the main production made for beauty production in pp collisions at LO.	66
5.2	Decay of $B \rightarrow D^* X$	67
5.3	Decay of $D^{*+} \rightarrow D^0 \pi_s^+$ and of $D^0 \rightarrow K^- \pi_s^+$. These mesons are prompt D mesons.	68
	(a) Decay of $D^{*+} \rightarrow D^0 \pi_s^+$	68
	(b) Decay of $D^0 \rightarrow K^- \pi_s^+$	68
5.4	Decay of $B \rightarrow D^* X \rightarrow D^0 \pi_s X \rightarrow K \pi \pi_s X$ ($X = \ell \nu_\ell$) at meson level.	69
5.5	$B \rightarrow D^*$ conversion table extracted by Pythia6.	71
5.6	Example for the covered phase space of the measurement of prompt D meson production at 5 TeV.	72
5.7	Differential cross sections for $D^{*\pm}$ mesons as a function of p_T for data (points) compared to the NLO QCD calculations of FONLL, POWHEG+PYTHIA, POWHEG+HERWIG, MC@NLO and GM-VFNS (histograms) from ATLAS[25].	72
5.8	Measurement of prompt $D^{*\pm}$ double differential cross section as a function of p_T in $ \eta $ bins from CMS [26].	73
5.9	Comparison of differential cross sections for $D^{*\pm}$ meson production from CMS [28] at $\sqrt{s} = 13$ TeV with ATLAS [25] at $\sqrt{s} = 7$ TeV.	75
	(a) Comparison of differential cross sections for $D^{*\pm}$ meson production from CMS (black circles, prompt [28]) at $\sqrt{s} = 13$ TeV with ATLAS (red squares, prompt + non-prompt) at $\sqrt{s} = 7$ TeV [25].	75
	(b) Differential cross sections for prompt $D^{*\pm}$ meson production at $\sqrt{s} = 13$ TeV from CMS [28].	75

5.10	Differential cross section in pp collisions at $\sqrt{s} = 5.02$ TeV for $B \rightarrow D^0$ from CMS [24].	76
5.11	p_T -differential production cross section of D^0 mesons with $ y < 0.5$ in pp collisions at $\sqrt{s} = 7$ TeV from ALICE [27].	77
5.12	Differential production cross sections for D mesons in the central rapidity range $ y < 0.5$ in pp collisions at 5 TeV from ALICE [30].	78
5.13	Measurements and predictions for the absolute prompt $D^{*\pm}$ cross-sections at $\sqrt{s} = 5$ TeV from LHCb [22].	79
6.1	A sketch of the decay $D^* \rightarrow D^0 \pi_s \rightarrow K \pi \pi_s$ and a cartoon for this decay. . .	91
6.2	Migration between the reconstructed and generated D^* for p_T , ϕ and y	94
6.3	Δm distributions for the signal extraction in the higher and lower p_T region for D^*	96
6.4	Overview of signal extraction for $D^* \rightarrow K \pi \pi_s$ in small bins in p_T for $ y < 2.5$ and $p_T > 1.5$ GeV.	97
6.5	Signal number of events and its uncertainty for $D^* \rightarrow K \pi \pi_s$ in bins in p_T and $ y $ for $ y < 2.5$ and $p_T > 1.5$ GeV.	98
6.6	Signal extraction in the higher and lower p_T region for prompt D^* in the matched MC.	98
6.7	Signal extraction in higher and lower p_T region for non-prompt D^* in the matched MC.	99
6.8	Efficiency calculation for $D^* \rightarrow K \pi \pi_s$	100
(a)	$N_{\text{reco \& matched}} D^* \rightarrow K \pi \pi_s$	100
(b)	$N_{\text{true}} D^* \rightarrow K \pi \pi_s$	100
(c)	Efficiency table for the reconstruction of $D^* \rightarrow K \pi \pi_s$	100
(d)	3D-Efficiency table for $D^* \rightarrow K \pi \pi_s$	100
6.9	Efficiency calculation for <i>prompt</i> $D^* \rightarrow K \pi \pi_s$	101
(a)	$N_{\text{reco \& matched}} \text{ prompt } D^* \rightarrow K \pi \pi_s$	101
(b)	$N_{\text{true}} \text{ prompt } D^* \rightarrow K \pi \pi_s$	101
(c)	Efficiency table for prompt $D^* \rightarrow K \pi \pi_s$	101
(d)	3D-Efficiency table for prompt $D^* \rightarrow K \pi \pi_s$ in MC	101
6.10	Efficiency calculation for <i>non – prompt</i> $D^* \rightarrow K \pi \pi_s$	102
(a)	$N_{\text{reco \& matched}} \text{ non-prompt } D^* \rightarrow K \pi \pi_s$ in MC	102
(b)	$N_{\text{true}} \text{ non-prompt } D^* \rightarrow K \pi \pi_s$	102
(c)	Efficiency table for non-prompt $D^* \rightarrow K \pi \pi_s$	102
(d)	3D-Efficiency table for non-prompt $D^* \rightarrow K \pi \pi_s$	102
6.11	Inclusive D^* cross sections as a function of p_T in the rapidity range $ y < 0.5$	104

6.12	D^* cross sections as a function of p_T in different rapidity ranges.	105
(a)	$0.5 < y < 1$	105
(b)	$1.0 < y < 1.5$	105
(c)	$1.5 < y < 2$	105
(d)	$2 < y < 2.5$	105
6.13	Inclusive D^* cross sections as a function of p_T in the different rapidity ranges from $0 < y < 2.5$	106
6.14	A sketch illustrating the dca for non-prompt and prompt $D^* \rightarrow D^0 \pi_s \rightarrow K \pi \pi_s$ production.	107
6.15	Example for dca distribution for prompt and non-prompt D^*	108
6.16	Separation of charm and beauty in D meson final states in higher and lower p_T region for $ y < 2.5$ with dca.	111
(a)	Separation of charm and beauty in D meson final states in higher region p_T region for $ y < 2.5$ with dca.	111
(b)	Separation of charm and beauty in D meson final states in lower region p_T region for $ y < 2.5$ with dca.	111
6.17	Correlation plots of the decay length significance and the dca from the D^0 for charm and beauty.	112
6.18	Separation of charm and beauty in D^* meson final states performed via the method using in higher and lower p_T region for $ y < 2.5$ the product of the decay length significance and the dca from the D^0	113
6.19	Prompt and non-prompt D^* cross sections as a function of p_T in different rapidity ranges determined using the parameter dca for the charm beauty separation.	118
(a)	$ y < 0.5$	118
(b)	$ y < 1$	118
(c)	$1.0 < y < 2$	118
(d)	$ y < 0.5, y < 1, 1 < y < 2$	118
6.20	Prompt and non-prompt D^* cross sections as a function of p_T in different rapidity ranges determined using the parameter $dca \cdot d\text{ISig}$ for the charm beauty separation.	119
(a)	$ y < 0.5$	119
(b)	$ y < 1$	119
(c)	$1.0 < y < 2$	119
(d)	$0 < y < 2$	119
A.1	An illustration of all weak modes.	127

A.2	Decay length D meson distribution.	128
A.3	Decay length significance D meson distribution.	128
A.4	J/Psi D meson distribution.	129
B.1	Signal extraction for $p_T : 1 - 2$ GeV and $ y < 2.5$	135
(a)	$ y < 0.5$, $p_T : 1 - 2$ GeV	135
(b)	$0.5 < y < 1$, $p_T : 1 - 2$ GeV	135
(c)	$1 < y < 1.5$, $p_T : 1 - 2$ GeV	135
(d)	$1.5 < y < 2$, $p_T : 1 - 2$ GeV	135
(e)	$2 < y < 2.5$, $p_T : 1 - 2$ GeV	135
B.2	Signal extraction for $p_T : 2 - 3$ GeV and $ y < 2.5$	136
(a)	$ y < 0.5$, $p_T : 2 - 3$ GeV	136
(b)	$0.5 < y < 1$, $p_T : 2 - 3$ GeV	136
(c)	$1 < y < 1.5$, $p_T : 2 - 3$ GeV	136
(d)	$1.5 < y < 2$, $p_T : 2 - 3$ GeV	136
(e)	$2 < y < 2.5$, $p_T : 2 - 3$ GeV	136
B.3	Signal extraction for $p_T : 3 - 4$ GeV and $ y < 2.5$	137
(a)	$ y < 0.5$, $p_T : 3 - 4$ GeV	137
(b)	$0.5 < y < 1$, $p_T : 3 - 4$ GeV	137
(c)	$1 < y < 1.5$, $p_T : 3 - 4$ GeV	137
(d)	$1.5 < y < 2$, $p_T : 3 - 4$ GeV	137
(e)	$2 < y < 2.5$, $p_T : 3 - 4$ GeV	137
B.4	Signal extraction for $p_T : 4 - 5$ GeV and $ y < 2.5$	138
(a)	$ y < 0.5$, $p_T : 4 - 5$ GeV	138
(b)	$0.5 < y < 1$, $p_T : 4 - 5$ GeV	138
(c)	$1 < y < 1.5$, $p_T : 4 - 5$ GeV	138
(d)	$1.5 < y < 2$, $p_T : 4 - 5$ GeV	138
(e)	$2 < y < 2.5$, $p_T : 4 - 5$ GeV	138
B.5	Signal extraction for $p_T : 5 - 6$ GeV and $ y < 2.5$	139
(a)	$ y < 0.5$, $p_T : 5 - 6$ GeV	139
(b)	$0.5 < y < 1$, $p_T : 5 - 6$ GeV	139
(c)	$1 < y < 1.5$, $p_T : 5 - 6$ GeV	139
(d)	$1.5 < y < 2$, $p_T : 5 - 6$ GeV	139
(e)	$2 < y < 2.5$, $p_T : 5 - 6$ GeV	139
B.6	Signal extraction for $p_T : 6 - 7$ GeV and $ y < 2.5$	140
(a)	$ y < 0.5$, $p_T : 6 - 7$ GeV	140
(b)	$0.5 < y < 1$, $p_T : 6 - 7$ GeV	140

(c)	$1 < y < 1.5, p_T : 6 - 7 \text{ GeV}$	140
(d)	$1.5 < y < 2, p_T : 6 - 7 \text{ GeV}$	140
(e)	$2 < y < 2.5, p_T : 6 - 7 \text{ GeV}$	140
B.7	Signal extraction for $p_T : 7 - 8 \text{ GeV}$ and $ y < 2.5$	141
(a)	$ y < 0.5, p_T : 7 - 8 \text{ GeV}$	141
(b)	$0.5 < y < 1, p_T : 7 - 8 \text{ GeV}$	141
(c)	$1 < y < 1.5, p_T : 7 - 8 \text{ GeV}$	141
(d)	$1.5 < y < 2, p_T : 7 - 8 \text{ GeV}$	141
(e)	$2 < y < 2.5, p_T : 7 - 8 \text{ GeV}$	141
B.8	Signal extraction for $p_T : 8 - 9 \text{ GeV}$ and $ y < 2.5$	142
(a)	$ y < 0.5, p_T : 8 - 9 \text{ GeV}$	142
(b)	$0.5 < y < 1, p_T : 8 - 9 \text{ GeV}$	142
(c)	$1 < y < 1.5, p_T : 8 - 9 \text{ GeV}$	142
(d)	$1.5 < y < 2, p_T : 8 - 9 \text{ GeV}$	142
(e)	$2 < y < 2.5, p_T : 8 - 9 \text{ GeV}$	142
B.9	Signal extraction for $p_T : 9 - 10 \text{ GeV}$ and $ y < 2.5$	143
(a)	$ y < 0.5, p_T : 9 - 10 \text{ GeV}$	143
(b)	$0.5 < y < 1, p_T : 9 - 10 \text{ GeV}$	143
(c)	$1 < y < 1.5, p_T : 9 - 10 \text{ GeV}$	143
(d)	$1.5 < y < 2, p_T : 9 - 10 \text{ GeV}$	143
(e)	$2 < y < 2.5, p_T : 9 - 10 \text{ GeV}$	143
B.10	Signal extraction for $p_T : > 10 \text{ GeV}$ and $ y < 2.5$	144
(a)	$ y < 0.5, p_T : > 10 \text{ GeV}$	144
(b)	$0.5 < y < 1, p_T : > 10 \text{ GeV}$	144
(c)	$1 < y < 1.5, p_T : > 10 \text{ GeV}$	144
(d)	$1.5 < y < 2, p_T : > 10 \text{ GeV}$	144
(e)	$2 < y < 2.5, p_T : > 10 \text{ GeV}$	144
C.1	Signal extraction for $p_T : 1 - 2 \text{ GeV}$ and $ y < 2$	148
(a)	$ y < 0.5, p_T : 1 - 2 \text{ GeV}$	148
(b)	$0.5 < y < 1, p_T : 1 - 2 \text{ GeV}$	148
(c)	$1 < y < 1.5, p_T : 1 - 2 \text{ GeV}$	148
C.2	Signal extraction for $p_T : 2 - 3 \text{ GeV}$ and $ y < 2$	149
(a)	$ y < 0.5, p_T : 2 - 3 \text{ GeV}$	149
(b)	$0.5 < y < 1, p_T : 2 - 3 \text{ GeV}$	149
(c)	$1 < y < 1.5, p_T : 2 - 3 \text{ GeV}$	149
C.3	Signal extraction for $p_T : 3 - 4 \text{ GeV}$ and $ y < 2$	150

(a)	$ y < 0.5, p_T : 3 - 4 \text{ GeV}$	150
(b)	$0.5 < y < 1, p_T : 3 - 4 \text{ GeV}$	150
(c)	$1 < y < 1.5, p_T : 3 - 4 \text{ GeV}$	150
C.4	Signal extraction for $p_T : 4 - 5 \text{ GeV}$ and $ y < 2$	151
(a)	$ y < 0.5, p_T : 4 - 5 \text{ GeV}$	151
(b)	$0.5 < y < 1, p_T : 4 - 5 \text{ GeV}$	151
(c)	$1 < y < 1.5, p_T : 4 - 5 \text{ GeV}$	151
C.5	Signal extraction for $p_T : 5 - 6 \text{ GeV}$ and $ y < 2$	152
(a)	$ y < 0.5, p_T : 5 - 6 \text{ GeV}$	152
(b)	$0.5 < y < 1, p_T : 5 - 6 \text{ GeV}$	152
(c)	$1 < y < 1.5, p_T : 5 - 6 \text{ GeV}$	152
C.6	Signal extraction for $p_T : 6 - 7 \text{ GeV}$ and $ y < 2$	153
(a)	$ y < 0.5, p_T : 6 - 7 \text{ GeV}$	153
(b)	$0.5 < y < 1, p_T : 6 - 7 \text{ GeV}$	153
(c)	$1 < y < 1.5, p_T : 6 - 7 \text{ GeV}$	153
C.7	Signal extraction for $p_T : 7 - 8 \text{ GeV}$ and $ y < 2$	154
(a)	$ y < 0.5, p_T : 7 - 8 \text{ GeV}$	154
(b)	$0.5 < y < 1, p_T : 7 - 8 \text{ GeV}$	154
(c)	$1 < y < 1.5, p_T : 7 - 8 \text{ GeV}$	154
C.8	Signal extraction for $p_T : 8 - 9 \text{ GeV}$ and $ y < 2$	155
(a)	$ y < 0.5, p_T : 8 - 9 \text{ GeV}$	155
(b)	$0.5 < y < 1, p_T : 8 - 9 \text{ GeV}$	155
(c)	$1 < y < 1.5, p_T : 8 - 9 \text{ GeV}$	155
C.9	Signal extraction for $p_T : 9 - 10 \text{ GeV}$ and $ y < 2$	156
(a)	$ y < 0.5, p_T : 9 - 10 \text{ GeV}$	156
(b)	$0.5 < y < 1, p_T : 9 - 10 \text{ GeV}$	156
(c)	$1 < y < 1.5, p_T : 9 - 10 \text{ GeV}$	156
C.10	Signal extraction for $p_T : > 10 \text{ GeV}$ and $ y < 2$	157
(a)	$ y < 0.5, p_T : > 10 \text{ GeV}$	157
(b)	$0.5 < y < 1, p_T : > 10 \text{ GeV}$	157
(c)	$1 < y < 1.5, p_T : > 10 \text{ GeV}$	157
C.11	Signal extraction in MC for $p_T : 1 - 2 \text{ GeV}$ and $ y < 2$	158
(a)	$ y < 0.5, p_T : 1 - 2 \text{ GeV}$	158
(b)	$0.5 < y < 1, p_T : 1 - 2 \text{ GeV}$	158
(c)	$1 < y < 1.5, p_T : 1 - 2 \text{ GeV}$	158
C.12	Signal extraction in MC for $p_T : 2 - 3 \text{ GeV}$ and $ y < 2$	159

(a)	$ y < 0.5, p_T : 2 - 3 \text{ GeV}$	159
(b)	$0.5 < y < 1, p_T : 2 - 3 \text{ GeV}$	159
(c)	$1 < y < 1.5, p_T : 2 - 3 \text{ GeV}$	159
C.13	Signal extraction in MC for $p_T : 3 - 4 \text{ GeV}$ and $ y < 2$	160
(a)	$ y < 0.5, p_T : 3 - 4 \text{ GeV}$	160
(b)	$0.5 < y < 1, p_T : 3 - 4 \text{ GeV}$	160
(c)	$1 < y < 1.5, p_T : 3 - 4 \text{ GeV}$	160
C.14	Signal extraction in MC for $p_T : 4 - 5 \text{ GeV}$ and $ y < 2$	161
(a)	$ y < 0.5, p_T : 4 - 5 \text{ GeV}$	161
(b)	$0.5 < y < 1, p_T : 4 - 5 \text{ GeV}$	161
(c)	$1 < y < 1.5, p_T : 4 - 5 \text{ GeV}$	161
C.15	Signal extraction in MC for $p_T : 5 - 6 \text{ GeV}$ and $ y < 2$	162
(a)	$ y < 0.5, p_T : 5 - 6 \text{ GeV}$	162
(b)	$0.5 < y < 1, p_T : 5 - 6 \text{ GeV}$	162
(c)	$1 < y < 1.5, p_T : 5 - 6 \text{ GeV}$	162
C.16	Signal extraction in MC for $p_T : 6 - 7 \text{ GeV}$ and $ y < 2$	163
(a)	$ y < 0.5, p_T : 6 - 7 \text{ GeV}$	163
(b)	$0.5 < y < 1, p_T : 6 - 7 \text{ GeV}$	163
(c)	$1 < y < 1.5, p_T : 6 - 7 \text{ GeV}$	163
C.17	Signal extraction in MC for $p_T : 7 - 8 \text{ GeV}$ and $ y < 2$	164
(a)	$ y < 0.5, p_T : 7 - 8 \text{ GeV}$	164
(b)	$0.5 < y < 1, p_T : 7 - 8 \text{ GeV}$	164
(c)	$1 < y < 1.5, p_T : 7 - 8 \text{ GeV}$	164
C.18	Signal extraction in MC for $p_T : 8 - 9 \text{ GeV}$ and $ y < 2$	165
(a)	$ y < 0.5, p_T : 8 - 9 \text{ GeV}$	165
(b)	$0.5 < y < 1, p_T : 8 - 9 \text{ GeV}$	165
(c)	$1 < y < 1.5, p_T : 8 - 9 \text{ GeV}$	165
C.19	Signal extraction in MC for $p_T : 9 - 10 \text{ GeV}$ and $ y < 2$	166
(a)	$ y < 0.5, p_T : 9 - 10 \text{ GeV}$	166
(b)	$0.5 < y < 1, p_T : 9 - 10 \text{ GeV}$	166
(c)	$1 < y < 1.5, p_T : 9 - 10 \text{ GeV}$	166
C.20	Signal extraction in MC for $p_T : > 10 \text{ GeV}$ and $ y < 2$	167
(a)	$ y < 0.5, p_T : > 10 \text{ GeV}$	167
(b)	$0.5 < y < 1, p_T : > 10 \text{ GeV}$	167
(c)	$1 < y < 1.5, p_T : > 10 \text{ GeV}$	167
C.21	Signal extraction in MC for $p_T : 1 - 2 \text{ GeV}$ and $ y < 2$	168

(a)	$ y < 0.5, p_T : 1 - 2 \text{ GeV}$	168
(b)	$0.5 < y < 1, p_T : 1 - 2 \text{ GeV}$	168
(c)	$1 < y < 1.5, p_T : 1 - 2 \text{ GeV}$	168
C.22	Signal extraction in MC for $p_T : 2 - 3 \text{ GeV}$ and $ y < 2$	169
(a)	$ y < 0.5, p_T : 2 - 3 \text{ GeV}$	169
(b)	$0.5 < y < 1, p_T : 2 - 3 \text{ GeV}$	169
(c)	$1 < y < 1.5, p_T : 2 - 3 \text{ GeV}$	169
C.23	Signal extraction in MC for $p_T : 3 - 4 \text{ GeV}$ and $ y < 2$	170
(a)	$ y < 0.5, p_T : 3 - 4 \text{ GeV}$	170
(b)	$0.5 < y < 1, p_T : 3 - 4 \text{ GeV}$	170
(c)	$1 < y < 1.5, p_T : 3 - 4 \text{ GeV}$	170
C.24	Signal extraction in MC for $p_T : 4 - 5 \text{ GeV}$ and $ y < 2$	171
(a)	$ y < 0.5, p_T : 4 - 5 \text{ GeV}$	171
(b)	$0.5 < y < 1, p_T : 4 - 5 \text{ GeV}$	171
(c)	$1 < y < 1.5, p_T : 4 - 5 \text{ GeV}$	171
C.25	Signal extraction in MC for $p_T : 5 - 6 \text{ GeV}$ and $ y < 2$	172
(a)	$ y < 0.5, p_T : 5 - 6 \text{ GeV}$	172
(b)	$0.5 < y < 1, p_T : 5 - 6 \text{ GeV}$	172
(c)	$1 < y < 1.5, p_T : 5 - 6 \text{ GeV}$	172
C.26	Signal extraction in MC for $p_T : 6 - 7 \text{ GeV}$ and $ y < 2$	173
(a)	$ y < 0.5, p_T : 6 - 7 \text{ GeV}$	173
(b)	$0.5 < y < 1, p_T : 6 - 7 \text{ GeV}$	173
(c)	$1 < y < 1.5, p_T : 6 - 7 \text{ GeV}$	173
C.27	Signal extraction in MC for $p_T : 7 - 8 \text{ GeV}$ and $ y < 2$	174
(a)	$ y < 0.5, p_T : 7 - 8 \text{ GeV}$	174
(b)	$0.5 < y < 1, p_T : 7 - 8 \text{ GeV}$	174
(c)	$1 < y < 1.5, p_T : 7 - 8 \text{ GeV}$	174
C.28	Signal extraction in MC for $p_T : 8 - 9 \text{ GeV}$ and $ y < 2$	175
(a)	$ y < 0.5, p_T : 8 - 9 \text{ GeV}$	175
(b)	$0.5 < y < 1, p_T : 8 - 9 \text{ GeV}$	175
(c)	$1 < y < 1.5, p_T : 8 - 9 \text{ GeV}$	175
C.29	Signal extraction in MC for $p_T : 9 - 10 \text{ GeV}$ and $ y < 2$	176
(a)	$ y < 0.5, p_T : 9 - 10 \text{ GeV}$	176
(b)	$0.5 < y < 1, p_T : 9 - 10 \text{ GeV}$	176
(c)	$1 < y < 1.5, p_T : 9 - 10 \text{ GeV}$	176
C.30	Signal extraction in MC for $p_T : > 10 \text{ GeV}$ and $ y < 2$	177

(a)	$ y < 0.5, p_T : > 10 \text{ GeV}$	177
(b)	$0.5 < y < 1, p_T : > 10 \text{ GeV}$	177
(c)	$1 < y < 1.5, p_T : > 10 \text{ GeV}$	177
C.31	Charm beauty separation based on dca for $p_T : 1 - 2 \text{ GeV}$ and $ y < 2$	178
(a)	$ y < 0.5, p_T : 1 - 2 \text{ GeV}$	178
(b)	$0.5 < y < 1, p_T : 1 - 2 \text{ GeV}$	178
(c)	$1 < y < 1.5, p_T : 1 - 2 \text{ GeV}$	178
C.32	Charm beauty separation based on dca for $p_T : 2 - 3 \text{ GeV}$ and $ y < 2$	179
(a)	$ y < 0.5, p_T : 2 - 3 \text{ GeV}$	179
(b)	$0.5 < y < 1, p_T : 2 - 3 \text{ GeV}$	179
(c)	$1 < y < 1.5, p_T : 2 - 3 \text{ GeV}$	179
C.33	Charm beauty separation based on dca for $p_T : 3 - 4 \text{ GeV}$ and $ y < 2$	180
(a)	$ y < 0.5, p_T : 3 - 4 \text{ GeV}$	180
(b)	$0.5 < y < 1, p_T : 3 - 4 \text{ GeV}$	180
(c)	$1 < y < 1.5, p_T : 3 - 4 \text{ GeV}$	180
C.34	Charm beauty separation based on dca for $p_T : 4 - 5 \text{ GeV}$ and $ y < 2$	181
(a)	$ y < 0.5, p_T : 4 - 5 \text{ GeV}$	181
(b)	$0.5 < y < 1, p_T : 4 - 5 \text{ GeV}$	181
(c)	$1 < y < 1.5, p_T : 4 - 5 \text{ GeV}$	181
C.35	Charm beauty separation based on dca for $p_T : 5 - 6 \text{ GeV}$ and $ y < 2$	182
(a)	$ y < 0.5, p_T : 5 - 6 \text{ GeV}$	182
(b)	$0.5 < y < 1, p_T : 5 - 6 \text{ GeV}$	182
(c)	$1 < y < 1.5, p_T : 5 - 6 \text{ GeV}$	182
C.36	Charm beauty separation based on dca for $p_T : 6 - 7 \text{ GeV}$ and $ y < 2$	183
(a)	$ y < 0.5, p_T : 6 - 7 \text{ GeV}$	183
(b)	$0.5 < y < 1, p_T : 6 - 7 \text{ GeV}$	183
(c)	$1 < y < 1.5, p_T : 6 - 7 \text{ GeV}$	183
C.37	Charm beauty separation based on dca for $p_T : 7 - 8 \text{ GeV}$ and $ y < 2$	184
(a)	$ y < 0.5, p_T : 7 - 8 \text{ GeV}$	184
(b)	$0.5 < y < 1, p_T : 7 - 8 \text{ GeV}$	184
(c)	$1 < y < 1.5, p_T : 7 - 8 \text{ GeV}$	184
C.38	Charm beauty separation based on dca for $p_T : 8 - 9 \text{ GeV}$ and $ y < 2$	185
(a)	$ y < 0.5, p_T : 8 - 9 \text{ GeV}$	185
(b)	$0.5 < y < 1, p_T : 8 - 9 \text{ GeV}$	185
(c)	$1 < y < 1.5, p_T : 8 - 9 \text{ GeV}$	185
C.39	Charm beauty separation based on dca for $p_T : 9 - 10 \text{ GeV}$ and $ y < 2$	186

(a)	$ y < 0.5, p_T : 9 - 10 \text{ GeV}$	186
(b)	$0.5 < y < 1, p_T : 9 - 10 \text{ GeV}$	186
(c)	$1 < y < 1.5, p_T : 9 - 10 \text{ GeV}$	186
C.40	Charm beauty separation based on dca for $p_T : > 10 \text{ GeV}$ and $ y < 2$	187
(a)	$ y < 0.5, p_T : > 10 \text{ GeV}$	187
(b)	$0.5 < y < 1, p_T : > 10 \text{ GeV}$	187
(c)	$1 < y < 1.5, p_T : > 10 \text{ GeV}$	187
C.41	Charm beauty separation based on $dca \cdot dl_{sig}$ for $p_T : 1 - 2 \text{ GeV}$ and $ y < 2$	188
(a)	$ y < 0.5, p_T : 1 - 2 \text{ GeV}$	188
(b)	$0.5 < y < 1, p_T : 1 - 2 \text{ GeV}$	188
(c)	$1 < y < 1.5, p_T : 1 - 2 \text{ GeV}$	188
C.42	Charm beauty separation based on $dca \cdot dl_{sig}$ for $p_T : 2 - 3 \text{ GeV}$ and $ y < 2$	189
(a)	$ y < 0.5, p_T : 2 - 3 \text{ GeV}$	189
(b)	$0.5 < y < 1, p_T : 2 - 3 \text{ GeV}$	189
(c)	$1 < y < 1.5, p_T : 2 - 3 \text{ GeV}$	189
C.43	Charm beauty separation based on $dca \cdot dl_{sig}$ for $p_T : 3 - 4 \text{ GeV}$ and $ y < 2$	190
(a)	$ y < 0.5, p_T : 3 - 4 \text{ GeV}$	190
(b)	$0.5 < y < 1, p_T : 3 - 4 \text{ GeV}$	190
(c)	$1 < y < 1.5, p_T : 3 - 4 \text{ GeV}$	190
C.44	Charm beauty separation based on $dca \cdot dl_{sig}$ for $p_T : 4 - 5 \text{ GeV}$ and $ y < 2$	191
(a)	$ y < 0.5, p_T : 4 - 5 \text{ GeV}$	191
(b)	$0.5 < y < 1, p_T : 4 - 5 \text{ GeV}$	191
(c)	$1 < y < 1.5, p_T : 4 - 5 \text{ GeV}$	191
C.45	Charm beauty separation based on $dca \cdot dl_{sig}$ for $p_T : 5 - 6 \text{ GeV}$ and $ y < 2$	192
(a)	$ y < 0.5, p_T : 5 - 6 \text{ GeV}$	192
(b)	$0.5 < y < 1, p_T : 5 - 6 \text{ GeV}$	192
(c)	$1 < y < 1.5, p_T : 5 - 6 \text{ GeV}$	192
C.46	Charm beauty separation based on $dca \cdot dl_{sig}$ for $p_T : 6 - 7 \text{ GeV}$ and $ y < 2$	193
(a)	$ y < 0.5, p_T : 6 - 7 \text{ GeV}$	193
(b)	$0.5 < y < 1, p_T : 6 - 7 \text{ GeV}$	193
(c)	$1 < y < 1.5, p_T : 6 - 7 \text{ GeV}$	193
C.47	Charm beauty separation based on $dca \cdot dl_{sig}$ for $p_T : 7 - 8 \text{ GeV}$ and $ y < 2$	194
(a)	$ y < 0.5, p_T : 7 - 8 \text{ GeV}$	194
(b)	$0.5 < y < 1, p_T : 7 - 8 \text{ GeV}$	194
(c)	$1 < y < 1.5, p_T : 7 - 8 \text{ GeV}$	194
C.48	Charm beauty separation based on $dca \cdot dl_{sig}$ for $p_T : 8 - 9 \text{ GeV}$ and $ y < 2$	195

(a)	$ y < 0.5, p_T : 8 - 9 \text{ GeV}$	195
(b)	$0.5 < y < 1, p_T : 8 - 9 \text{ GeV}$	195
(c)	$1 < y < 1.5, p_T : 8 - 9 \text{ GeV}$	195
C.49	Charm beauty separation based on $dca \cdot dl_{sig}$ for $p_T : 9 - 10 \text{ GeV}$ and $ y < 2$	196
(a)	$ y < 0.5, p_T : 9 - 10 \text{ GeV}$	196
(b)	$0.5 < y < 1, p_T : 9 - 10 \text{ GeV}$	196
(c)	$1 < y < 1.5, p_T : 9 - 10 \text{ GeV}$	196
C.50	Charm beauty separation based on $dca \cdot dl_{sig}$ for $p_T : > 10 \text{ GeV}$ and $ y < 2$	197
(a)	$ y < 0.5, p_T : > 10 \text{ GeV}$	197
(b)	$0.5 < y < 1, p_T : > 10 \text{ GeV}$	197
(c)	$1 < y < 1.5, p_T : > 10 \text{ GeV}$	197
C.51	Correlation plots of the decay length and the dca from the D^0 for charm and beauty.	198
C.52	Correlation plots of the angle ϕ and the dca from the D^0 for charm and beauty.	198
C.53	Correlation plots of the decay length significance and the angle ϕ from the D^0 for charm and beauty.	199

List of tables

2.1	List of hypercharge values for left and right handed fermions.	12
5.1	Values for the fragmentation fractions $f(c \rightarrow \mathcal{H}_c)$ and $f(b \rightarrow \mathcal{H}_c)$ are derived from the LEP measurements [1].	66
5.2	List of properties of charm and beauty hadrons from PDG 2020 [5]. The mass, the quark content and the decay length of the individual hadron is shown.	67
5.3	List of branching fractions taken from PDG 2020 [5].	68
6.1	This table lists the different data samples used in this work. The list of 2015E 5 TeV pp datasets shows the name of the dedicated sample, the number of events without using a JSON file, and in the next column the next number with using it. The last column represents the minimum bias collisions, i.e. here the number of reconstructed good quality primary vertices (Section 6.2). The integrated luminosity is shown in the last row.	87
6.2	This table lists the JSON files of 2015 5 TeV proton proton collision. The symbol (*) represents the one used by CMS-HIN [2, 3] (Heavy Ion working group within CMS with which results are compared within this study) while the symbol (**) stands for the latest JSON file that is used within this work leading to more statistics compared to (*).	87
6.3	List of 2015 5 TeV pp MC generated through CMS-BPH (B Physics working group within CMS), full phase space. The table shows the name of the MC sample in the first column, its number of events in the second and in the last column the effective luminosity.	89
6.4	Preselection at AOD level.	92
6.5	Cuts on the decay products of D^*	93
6.6	List of criteria for the selection of D^* candidates.	93

- 6.7 Measured partial total cross sections in μb for prompt and non-prompt D^* production at 5 TeV in bins of p_T and in the rapidity range $|y| < 0.5$. The first bin is measured starting from 1.5 GeV and the last bin correspond to the overflow bin. All other bins have a bin width of 1 GeV. The charm beauty separation is performed with dca 117
- 6.8 Measured partial total cross sections in μb for prompt and non-prompt D^* production at 5 TeV in bins of p_T and in the rapidity range $|y| < 1$. The first bin is measured starting from 1.5 GeV and the last bin correspond to the overflow bin. All bins have a bin width from 1 GeV. The charm beauty separation is performed with dca 120
- 6.9 Measured partial total cross sections in μb for prompt and non-prompt D^* production at 5 TeV in bins of p_T and in the rapidity range $1 < |y| < 2$. The first bin is measured starting from 1.5 GeV and the last bin correspond to the overflow bin. All bins have a bin width from 1 GeV. The charm beauty separation is performed with dca 120
- 6.10 Measured partial total cross sections in μb for prompt and non-prompt D^* production at 5 TeV in bins of p_T and in the rapidity range $|y| < 0.5$. The first bin is measured starting from 1.5 GeV and the last bin correspond to the overflow bin. All bins have a bin width from 1 GeV. The charm beauty separation is performed with $dca \cdot dl_{sig}$ 121
- 6.11 Measured partial total cross sections in μb for prompt and non-prompt D^* production at 5 TeV in bins of p_T and in the rapidity range $|y| < 1$. The first bin is measured starting from 1.5 GeV and the last bin correspond to the overflow bin. All bins have a bin width from 1 GeV. The charm beauty separation is performed with $dca \cdot dl_{sig}$ 121
- 6.12 Measured partial total cross sections in μb for prompt and non-prompt D^* production at 5 TeV in bins of p_T and in the rapidity range $1 < |y| < 2$. The first bin is measured starting from 1.5 GeV and the last bin correspond to the overflow bin. All bins have a bin width from 1 GeV. The charm beauty separation is performed with $dca \cdot dl_{sig}$ 122